

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 10:53:38 ; Search time 6264.22 Seconds
(without alignments)
11470.669 Million cell updates/sec

Title: US-09-610-313-30

Perfect score: 2469

Sequence: 1 gtcgacgcacacatggccga.....gggctagcacgggtgaattc 2469

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_bt:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_on:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rod:*

36: em_htg_man:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	2469	100.0	2469	6	AX455914	Sequence
2	2442.2	98.9	2463	6	AX455915	Sequence
3	2415.4	97.8	2457	6	AX455916	Sequence
4	2051.2	83.1	9166	6	AX427930	Sequence
5	2042.8	82.7	9788	6	AX427936	Sequence
6	2028.8	82.2	9169	6	AX427931	Sequence
7	2027	82.1	9194	6	AX427925	Sequence
8	2027	82.1	12411	6	AX427927	Sequence
9	2025.4	82.0	9194	6	AX427926	Sequence
10	2020.4	81.8	9785	6	AX427938	Sequence
11	2018.6	81.8	9189	6	AX427921	Sequence
12	2017	81.7	9167	6	AX427933	Sequence
13	2017	81.7	9170	6	AX427928	Sequence
14	2017	81.7	9782	6	AX427935	Sequence
15	2017	81.7	9783	6	AX427934	Sequence
16	2017	81.7	9792	6	AX427932	Sequence
17	2015.6	81.6	4353	12	AF287353	Synthetic
18	1999	81.0	3015	6	AX455946	Sequence
19	1982.4	80.3	3009	6	AX455987	Sequence
20	1958.8	79.3	4352	12	AF287352	Synthetic
21	1907	77.2	2577	6	AX457088	Sequence
22	1905.8	77.2	4343	6	AX306428	Sequence
23	1901.8	77.0	4341	6	AX306429	Sequence
24	1892.2	76.6	1977	6	AX455954	Sequence
25	1879.2	76.1	7897	6	AX427923	Sequence
26	1837	74.4	1965	6	AX455952	Sequence
27	1825.6	73.9	3024	12	AF287355	Synthetic
28	1807.6	73.2	9407	6	AX427937	Sequence
29	1678.6	68.0	4307	6	AX019132	Sequence
30	1659.4	67.2	4307	6	AX035453	Sequence
31	1659.4	67.2	4307	6	AX056836	Sequence
32	1651.4	66.9	4307	6	AX283597	Sequence
33	1651.4	66.9	4307	6	AX283607	Sequence
34	1651.4	66.9	4327	6	AX035465	Sequence
35	1651.4	66.9	4353	6	AX035464	Sequence
36	1651.4	66.9	4642	6	AX035463	Sequence
37	1651.4	66.9	9772	6	AX283584	Sequence
38	1608	65.1	1680	6	AX455960	Sequence
39	1592.8	64.5	8366	6	AX180752	Sequence
40	1557.4	63.1	4338	6	AX180747	Sequence
41	1554.4	63.0	1668	6	AX455962	Sequence
42	1517.4	61.5	3981	6	AX306430	Sequence
43	1303.6	52.8	4288	6	AX149648	Sequence
44	1279.4	51.8	9010	14	AF110974	HIV-1 iso
45	1279.4	51.8	9010	14	AF110975	HIV-1 iso

ALIGNMENTS

RESULT 1	AX455914	Sequence 30 from Patent WO0204493.	2469 bp	DNA	linear	PAT 06-JUL-2002
LOCUS	AX455914					
DEFINITION	AX455914					
ACCESSION	AX455914					
VERSION	AX455914.1	GI:21714899				
KEYWORDS	.					
SOURCE	synthetic construct.					
ORGANISM	synthetic construct					
	artificial sequences.					
REFERENCE	1					
AUTHORS	zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.					
TITLE	Polynucleotides encoding antigenic hiv type c polypeptides,					
	polypeptides and uses thereof					
JOURNAL	Patent: WO 0204493-A 30 17-JAN-2002;					

Db 1981 GCCGAGACCTTTCTAGTGGAGCGCGCCCAACCCGAGAGCAAGATCGGCAAGGCCGCG 2040
Qy 2041 TACGTGACCGACCGCGCGCGCAGAAAGATCGTAGCCCTGACCGAGACCAACCAAGAG 2100
Db 2041 TACGTGACCGACCGCGCGCGCAGAAAGATCGTAGCCCTGACCGAGACCAACCAAGAG 2100
Qy 2101 ACCGAGCTGCAGGCCATCCAGCTGSCCTCGAGGACAGCGGAGAGGTGAACATCGTG 2160
Db 2101 ACCGAGCTGCAGGCCATCCAGCTGSCCTCGAGGACAGCGGAGAGGTGAACATCGTG 2160
Qy 2161 ACCGACACCGAGTACCGCTTGGGATCATCATGAGGCCACCGCGACAGAGGAGCGAG 2220
Db 2161 ACCGACACCGAGTACCGCTTGGGATCATCATGAGGCCACCGCGACAGAGGAGCGAG 2220
Qy 2221 CTGTTGAACACAGATCATCGACGAGCTGATCAAGAGGAGAGGTGATCTGAGCTGGGTG 2280
Db 2221 CTGTTGAACACAGATCATCGACGAGCTGATCAAGAGGAGAGGTGATCTGAGCTGGGTG 2280
Qy 2281 CCGGCCCAACAGGGGATCGCGGCAAGAGCAGATCGACAGCTGAGCAAGGGCATC 2340
Db 2281 CCGGCCCAACAGGGGATCGCGGCAAGAGCAGATCGACAGCTGAGCAAGGGCATC 2340
Qy 2341 CGCAAGTGTCTTCTTGGAGGGATCATCGATGGCGCATCTGATCTACCATGATGAC 2400
Db 2341 CGCAAGTGTCTTCTTGGAGGGATCATCGATGGCGCATCTGATCTACCATGATGAC 2400
Qy 2401 GACCTGTACGTGGCGAGCGCGGCTAGGATCGATTAAAGCTTCCCGGGGTAGCAC 2460
Db 2401 GACCTGTACGTGGCGAGCGCGGCTAGGATCGATTAAAGCTTCCCGGGGTAGCAC 2460
Qy 2461 GGTGAATTC 2469
Db 2461 GGTGAATTC 2469

RESULT 2
AX455915
LOCUS AX455915 2463 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 31 from Patent WO0204493.
ACCESSION AX455915
VERSION AX455915.1 GI:21714900

SOURCE
ORGANISM
synthetic construct.
synthetic construct
artificial sequences.

REFERENCE
1

AUTHORS
TITLE zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.
Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof

JOURNAL Patent: WO 0204493-A 31 17-JAN-2002;

CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

FEATURES
Location/Qualifiers

Source 1. .2463

/organism="synthetic construct"

/db.xref="taxon:32630"

/note="PR975YM"

BASE COUNT 567 a 835 c 759 g 302 t

ORIGIN

Query Match 98.9%; Score 2442.2; DB 6; Length 2463;

Best Local Similarity 99.6%; Pred. No. 8.3e-248;

Matches 2460; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

Qy 1 GTCGACGCCACATGGCGGAGGCCATGAGCGAGGCCACCGAGCGCAACATCCTGATGCAG 60

Db 1 GTCGACGCCACATGGCGGAGGCCATGAGCGAGGCCACCGAGCGCAACATCCTGATGCAG 60

Qy 61 CGCAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTCTTCAACTCGCGCAAGGAGGCG 120

Db 61 CGCAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTCTTCAACTCGCGCAAGGAGGCG 120

Qy 121 CACATCGCCCGCACTGCGCGGCCCGCCGCAAGAGGGTCTGCTGAAGTGGCGCAAGGAG 180

Db 121 CACATCGCCCGCACTGCGCGGCCCGCCGCAAGAGGGTCTGCTGAAGTGGCGCAAGGAG 180
Qy 181 GGCACACAGATGAAGAGTATGACCGAGCGCCAGGCCCAACTTCTTCGGGAGAGACCTGGCC 240
Db 181 GGCACACAGATGAAGAGTATGACCGAGCGCCAGGCCCAACTTCTTCGGGAGAGACCTGGCC 240
Qy 241 TTCCCGCCAGGGCAAGCGCGGAGTTCCCGAGCGAGCAGAACCGCGCCCAACAGCCCGCAC 300
Db 241 TTCCCGCCAGGGCAAGCGCGGAGTTCCCGAGCGAGCAGAACCGCGCCCAACAGCCCGCAC 300
Qy 301 AGCCCGAGCTGACAGTGCAGGTGCGCGGCGACAAACCCCGCAGCGAGCGCGCGCGCGCCAG 360
Db 301 AGCCCGAGCTGACAGTGCAGGTGCGCGGCGACAAACCCCGCAGCGAGCGCGCGCGCGCCAG 360
Qy 361 GGCACCTTGAACCTTCCCGCAGATCACCTGTGGCAGCGCCCTGTGTGAGCATCAAGGTG 420
Db 361 GGCACCTTGAACCTTCCCGCAGATCACCTGTGGCAGCGCCCTGTGTGAGCATCAAGGTG 420
Qy 421 GCGGCCAGATCAAGGAGGCCCTGTGGACACCGCGCGCGAGCACACCGTGTGGAGGAG 480
Db 421 GCGGCCAGATCAAGGAGGCCCTGTGTGGACACCGCGCGCGAGCACACCGTGTGGAGGAG 480
Qy 481 ATGAGCTTCCCGGCAAGTGGAGCCCAAGATGATCGCGGCATCGCGGCTTTCATCAAG 540
Db 481 ATGAGCTTCCCGGCAAGTGGAGCCCAAGATGATCGCGGCATCGCGGCTTTCATCAAG 540
Qy 541 GTGCGCCAGTACGACACAGATCTGTAGAGATCTCGGCAAGAGGCATCGGCAAGGAGG 600
Db 541 GTGCGCCAGTACGACACAGATCTGTAGAGATCTCGGCAAGAGGCATCGGCAAGGAGG 600
Qy 601 CTGATCGGCCCGCCACCGCGTGAACATCATCGGCCCAACATCTGACCCAGCTGAGCGGTG 660
Db 601 CTGATCGGCCCGCCACCGCGTGAACATCATCGGCCCAACATCTGACCCAGCTGAGCGGTG 660
Qy 661 ACCCTGAACCTTCCCGCATCGACCCCATCGAGACCGTGCCTGAAAGCTGAAGCCCGGCATG 720
Db 661 ACCCTGAACCTTCCCGCATCGACCCCATCGAGACCGTGCCTGAAAGCTGAAGCCCGGCATG 720
Qy 721 GACGCCCCAAGGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGGCTTGCACCGCC 780
Db 721 GACGCCCCAAGGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGGCTTGCACCGCC 780
Qy 781 ATCTCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTAC 840
Db 781 ATCTCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTAC 840
Qy 841 AACACCCCGCTTTCGCCCATCAAGAAAGAGACACCAAGTGGCGCAAGCTGGTGGAC 900
Db 841 AACACCCCGCTTTCGCCCATCAAGAAAGAGACACCAAGTGGCGCAAGCTGGTGGAC 900
Qy 901 TTCCGCGAGCTGAACAAGCGCACCGACACTTCTGGGAGGTGCAGCTGGGCATCCCGCAC 960
Db 901 TTCCGCGAGCTGAACAAGCGCACCGACACTTCTGGGAGGTGCAGCTGGGCATCCCGCAC 960
Qy 961 CCGCGCGCCCTGAAGAAGAAGAGCGTGCCTGCTGGAGTGGCGAGCGCTTACCTTC 1020
Db 961 CCGCGCGCCCTGAAGAAGAAGAGCGTGCCTGCTGGAGTGGCGAGCGCTTACCTTC 1020
Qy 1021 AGCGTGCCTTGGAGGAGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCAAC 1080
Db 1021 AGCGTGCCTTGGAGGAGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCAAC 1080
Qy 1081 AACGAGACCCCGCATCCGCTACAGTACACGCTGCTGGAGCGCTTCCGCGCGCGCAAGG 1140
Db 1081 AACGAGACCCCGCATCCGCTACAGTACACGCTGCTGGAGCGCTTCCGCGCGCGCAAGG 1140
Qy 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCGCAAGG 1200
Db 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCGCAAGG 1200
Qy 1201 GAGATCGTGTACTACAGTACATGGAGCACTGTACGTGGGAGCGAGCTTGAGATCGGC 1260
Db 1201 GAGATCGTGTACTACAGTACATGGAGCACTGTACGTGGGAGCGAGCTTGAGATCGGC 1260

Db 1201 GAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGCGAGCAGCCTGGAGATCGGC 1254
QY 1261 CAGCACCGCGCAAGATCGAGAGCTGCGCAAGCACTCTGCTGCGTGGGCTTCACCAACC 1320
Db 1255 CAGCACCGCGCAAGATCGAGAGCTGCGCAAGCACTCTGCTGCGTGGGCTTCACCAACC 1314
QY 1321 CCGACAAGAAGCACCAAGAGAGCCCTTCTCTGTGATGGGCTACAGCTGCAACCC 1380
Db 1315 CCGACAAGAAGCACCAAGAGAGCCCTTCTCTGTGATGGGCTACAGCTGCAACCC 1374
QY 1381 GACAAGTGGACCGTGCAGCCATCGAGCTGCCGAGAGAGAGCTGGACCTGGAACGAC 1440
Db 1375 GACAAGTGGACCGTGCAGCCATCGAGCTGCCGAGAGAGAGCTGGACCTGGAACGAC 1434
QY 1441 ATCCAGAAGCTGTGGCGAAGCTGAACTGGCGCAGCAGATCTACCCGCGCATCAAGGTG 1500
Db 1435 ATCCAGAAGCTGTGGCGAAGCTGAACTGGCGCAGCAGATCTACCCGCGCATCAAGGTG 1494
QY 1501 CGCCAGCTGTGAAGTGTGCGCGCGCCAGAGCCCTGACCCACATCTGTCGCCCTGACC 1560
Db 1495 CGCCAGCTGTGAAGTGTGCGCGCGCCAGAGCCCTGACCCACATCTGTCGCCCTGACC 1554
QY 1561 GAGGAGCGAGCTGAGCTGGCCGAGAGACCGGAGATCTTGGCGAGCCCTGCGACGGC 1620
Db 1555 GAGGAGCGAGCTGAGCTGGCCGAGAGACCGGAGATCTTGGCGAGCCCTGCGACGGC 1614
QY 1621 GTGTACTAGCACCCAGCAAGGACCTGTGGCGGAGATCCAGAGCAGGCGCACGACCCAG 1680
Db 1615 GTGTACTAGCACCCAGCAAGGACCTGTGGCGGAGATCCAGAGCAGGCGCACGACCCAG 1674
QY 1681 TGGACCTACCAAGTCTACCAAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAG 1740
Db 1675 TGGACCTACCAAGTCTACCAAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAG 1734
QY 1741 ATCCGACCGCCACACCAAGGACCTGAGACGACTACCGAGGCGCTGAGAGATCGCC 1800
Db 1735 ATCCGACCGCCACACCAAGGACCTGAGACGACTACCGAGGCGCTGAGAGATCGCC 1794
QY 1801 ATGAGAGATCGTGATCTGGGCAAGACCCCAAGCTTCGCGCTGCCCATCCAGAGGAG 1860
Db 1795 ATGAGAGATCGTGATCTGGGCAAGACCCCAAGCTTCGCGCTGCCCATCCAGAGGAG 1854
QY 1861 ACCTGGAGACCTGTGGACCGACTACTGGCAGGCGACCTTGGATCCCCGAGTGGGAGTTC 1920
Db 1855 ACCTGGAGACCTGTGGACCGACTACTGGCAGGCGACCTTGGATCCCCGAGTGGGAGTTC 1914
QY 1921 GTCAACACCCCTGCTGTGAGCTGTGTACAGCTGGAGAGGAGCCCATCATCTCGC 1980
Db 1915 GTCAACACCCCTGCTGTGAGCTGTGTACAGCTGGAGAGGAGCCCATCATCTCGC 1974
QY 1981 GCGAGACCTTCTACGTGGACGGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCGCGC 2040
Db 1975 GCGAGACCTTCTACGTGGACGGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCGCGC 2034
QY 2041 TAGCTGACCGACCGGGCGCGGAGAGATGTGAGGCTTACCGAGACCAACCAACCAAGAG 2100
Db 2035 TAGCTGACCGACCGGGCGCGGAGAGATGTGAGGCTTACCGAGACCAACCAACCAAGAG 2094
QY 2101 ACCGAGCTGACGCGCATCCAGCTGGCCCTGCAGGACAGCGGCGAGCGGTGAACATCGTG 2160
Db 2095 ACCGAGCTGACGCGCATCCAGCTGGCCCTGCAGGACAGCGGCGAGCGGTGAACATCGTG 2154
QY 2161 ACCGAGCGAGTACGCGCTTGGGCATCATCCAGGCGCCAGCCGACAGAGCGAGCGAG 2220
Db 2155 ACCGAGCGAGTACGCGCTTGGGCATCATCCAGGCGCCAGCCGACAGAGCGAGCGAG 2214
QY 2221 CTGGTGAACAGATCATCGAGCAGCTGATCAAGAGAGAGGTGTACCTGAGCTGGGTG 2280
Db 2215 CTGGTGAACAGATCATCGAGCAGCTGATCAAGAGAGAGGTGTACCTGAGCTGGGTG 2274
QY 2281 CCGCGCCACAGGCGATCGCGCGCACGAGCAGATCGACAAAGCTGGTGAAGGCGCATC 2340
Db 2275 CCGCGCCACAGGCGATCGCGCGCACGAGCAGATCGACAAAGCTGGTGAAGGCGCATC 2334

QY 2341 CGCAAGTGTCTGTTCTGGACGGCATCGATGGCGGATCGTGTATCTACCAAGTACATGGAC 2400
Db 2335 CGCAAGTGTCTGTTCTGGACGGCATCGATGGCGGATCGTGTATCTACCAAGTACATGGAC 2394
QY 2401 GACCTGTAGTGGGAGCGGCGCCCTAGATCGATTAAAGCTTCCCGGGGTAGCACC 2460
Db 2395 GACCTGTAGTGGGAGCGGCGCCCTAGATCGATTAAAGCTTCCCGGGGTAGCACC 2454
QY 2461 GGTGAATTC 2469
Db 2455 GGTGAATTC 2463
RESULT 3
AX455916
LOCUS AX455916
DEFINITION Sequence 32 from Patent WO0204493.
ACCESSION AX455916
VERSION AX455916.1 GI:21714901
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL Polypeptides and uses thereof
PUBLISHED WO 0204493-A 32 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
Location/Qualifiers
1..2457
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="PR975YMHM"
BASE COUNT 566 a 837 c 754 g 300 t
ORIGIN
Query Match 97.8%; Score 2415.4; DB 6; Length 2457;
Best Local Similarity 99.3%; Pred. No. 5.4e-245;
Matches 2451; Conservative 0; Mismatches 6; Indels 12; Gaps 2;
QY 1 GTCGAGCCACCATGCGGAGGCCATGAGCCAGGCCACAGCCGCAACATCTCTGATGCAG 60
Db 1 GTCGAGCCACCATGCGGAGGCCATGAGCCAGGCCACAGCCGCAACATCTCTGATGCAG 60
QY 61 CGCAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGC 120
Db 61 CGCAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGC 120
QY 121 CACATCGCCCGCAACTGCGCGCGCCCGCCGCAAGAAGGGCTGCTGGAAGTGCAGGAG 180
Db 121 CACATCGCCCGCAACTGCGCGCGCCCGCCGCAAGAAGGGCTGCTGGAAGTGCAGGAG 180
QY 181 GGCACACAGATGAAGGACTGCACCGGAGCCGAGGCCAACTTCTTCGCGAGGACCTGGCC 240
Db 181 GGCACACAGATGAAGGACTGCACCGGAGCCGAGGCCAACTTCTTCGCGAGGACCTGGCC 240
QY 241 TTCCCCCGAGGCAAGCCCGCGAGTTCCCGCAGCAGAGAACCGCCCAACAGCCCCCACC 300
Db 241 TTCCCCCGAGGCAAGCCCGCGAGTTCCCGCAGCAGAGAACCGCCCAACAGCCCCCACC 300
QY 301 AGCGCGAGCTGCAAGTTCGCGGCGGACAAACCCCGCAGCGAGCCGCGCGCGAGCGCAG 360
Db 301 AGCGCGAGCTGCAAGTTCGCGGCGGACAAACCCCGCAGCGAGCCGCGCGCGAGCGCAG 360
QY 361 GGCACCCCTGAACCTTCCCGCAGATCACCTGTGGAGCGCCCTCTGTGAGCATCAAGGTG 420
Db 361 GGCACCCCTGAACCTTCCCGCAGATCACCTGTGGAGCGCCCTCTGTGAGCATCAAGGTG 420
QY 421 GCGCGCCAGATCAAGGAGGCCCTGTGGACACCGCGCCGACACACCGTGTGAGGAG 480
Db 421 GCGCGCCAGATCAAGGAGGCCCTGTGGACACCGCGCCGACACACCGTGTGAGGAG 480

QY 481 ATGAGCCTGCGCGCAAGTGGAGCCCAAGATGATCGCGGCATCGCGGCTTTCATCAAG 540
DB 481 ATGAGCCTGCGCGCAAGTGGAGCCCAAGATGATCGCGGCATCGCGGCTTTCATCAAG 540
QY 541 GTGCGCAGTAGCAGACAGATCTCTGATCGAGATCTGCGGCAAGAGCCCATCGGACCCGTG 600
DB 541 GTGCGCAGTAGCAGACAGATCTCTGATCGAGATCTGCGGCAAGAGCCCATCGGACCCGTG 600
QY 601 CTGATCGGCCCCACCCCGGTGAACATCATCGGCGCAAGATGCTGACCCAGCTGGGCTGC 660
DB 601 CTGATCGGCCCCACCCCGGTGAACATCATCGGCGCAAGATGCTGACCCAGCTGGGCTGC 660
QY 661 ACCCTGAATCTCCCATCATGAGCCCATCGAGACGTCGCGTGAAGCTGAAGCCCGCATG 720
DB 661 ACCCTGAATCTCCCATCATGAGCCCATCGAGACGTCGCGTGAAGCTGAAGCCCGCATG 720
QY 721 GACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAATCAAGGCCCTGACCGCC 780
DB 721 GACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAATCAAGGCCCTGACCGCC 780
QY 781 ATCTCGAGGAGTAGAGAGGAGGCAAGATCACCAGATCGGCCCCCGAGAACCCCTAC 840
DB 781 ATCTCGAGGAGTAGAGAGGAGGCAAGATCACCAGATCGGCCCCCGAGAACCCCTAC 840
QY 841 AACACCCCGTGTTCGCCATCAAGAAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
DB 841 AACACCCCGTGTTCGCCATCAAGAAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
QY 901 TTCGCGAGCTGAACAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCAATCCCCAC 960
DB 901 TTCGCGAGCTGAACAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCAATCCCCAC 960
QY 961 CCGCGCGGCTGAAGAAAGAGAGCGTGCACGCTGAGCTGGGCGAGCGCTACTTC 1020
DB 961 CCGCGCGGCTGAAGAAAGAGAGCGTGCACGCTGAGCTGGGCGAGCGCTACTTC 1020
QY 1021 AGCGTCCCTTGACGAGGACTTCCGCAAGTACACGCGCTTACCATCCCGCAGCATCAAC 1080
DB 1021 AGCGTCCCTTGACGAGGACTTCCGCAAGTACACGCGCTTACCATCCCGCAGCATCAAC 1080
QY 1081 AAGCAGACCCCGGCTACCGCTACCAAGTACAGCTGTCGCGCCAGGGCTGGAAGGCGAG 1140
DB 1081 AAGCAGACCCCGGCTACCGCTACCAAGTACAGCTGTCGCGCCAGGGCTGGAAGGCGAG 1140
QY 1141 CCGAGATCTTCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACCCC 1200
DB 1141 CCGAGATCTTCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACCCC 1200
QY 1201 GAGATCGTGATCTACCATGATGAGAGCTGTAGCTGGGCGAGCAGCTGGAGATCGGC 1260
DB 1201 GAGATCGTGATCTACCA - - - - -GGCCCCCTGTACGTGGGCGAGCAGCTGGAGATCGGC 1264
QY 1261 CAGCAGCGGCGCAGATCGAGGAGCTGCGCAAGCAGCTGTCGCGCTGGGCGCTTACCAACC 1320
DB 1255 CAGCAGCGGCGCAGATCGAGGAGCTGCGCAAGCAGCTGTCGCGCTGGGCGCTTACCAACC 1314
QY 1321 CCGCACAAGAGCAGCAGAGGAGCCCTTCTCTGTGTGATGGCTACGAGCTGCACCCC 1380
DB 1315 CCGCACAAGAGCAGCAGAGGAGCCCTTCTCTGTGTGATGGCTACGAGCTGCACCCC 1368
QY 1381 GACAAGTGAACCTGACGCCATCGAGCTGCCGGAAGAGAGCTGGACCGGTGAAGCAG 1440
DB 1369 GACAAGTGAACCTGACGCCATCGAGCTGCCGGAAGAGAGCTGGACCGGTGAAGCAG 1428
QY 1441 ATCCAGAAGCTGTGGGCAAGCTGAATCTGGGCGCAGCAGATCTACCCCGCATCAAGGTG 1500
DB 1429 ATCCAGAAGCTGTGGGCAAGCTGAATCTGGGCGCAGCAGATCTACCCCGCATCAAGGTG 1488
QY 1501 CGCCAGCTGTCAAGCTGTGCGGGCGCAGGCGCTGACGACATCTGTCCTTACCC 1560
DB 1489 CGCCAGCTGTCAAGCTGTGCGGGCGCAGGCGCTGACGACATCTGTCCTTACCC 1548

QY 1561 GAGGAGGCGAGCTGGAGCTGGCCGAGAACCCGAGATCTGCGCGAGCCGCTGCACGCG 1620
DB 1549 GAGGAGGCGAGCTGGAGCTGGCCGAGAACCCGAGATCTGCGCGAGCCGCTGCACGCG 1608
QY 1621 GTGTACTAGACACCCAGCAGAGACCTGGTGGCGAGATCCAGAAAGAGGCGCCACGACGAG 1680
DB 1609 GTGTACTAGACACCCAGCAGAGACCTGGTGGCGAGATCCAGAAAGAGGCGCCACGACGAG 1668
QY 1681 TGACCTACCCAGATCTACCCAGGAGCCCTTCAAGAACCTGAGACCCGCAAGTACGCCAAG 1740
DB 1669 TGACCTACCCAGATCTACCCAGGAGCCCTTCAAGAACCTGAGACCCGCAAGTACGCCAAG 1728
QY 1741 ATCGCAGCCGCCACACCAACACGCTGAAGCAGCTGACCCAGGCGCTGCAAGAAGATCGCC 1800
DB 1729 ATCGCAGCCGCCACACCAACACGCTGAAGCAGCTGACCCAGGCGCTGCAAGAAGATCGCC 1788
QY 1801 ATGAGAGCATCTGTATCTGGGCGAAGACCCCAAGTTCGCTGCCCTCAGAAAGGAG 1860
DB 1789 ATGAGAGCATCTGTATCTGGGCGAAGACCCCAAGTTCGCTGCCCTCAGAAAGGAG 1848
QY 1861 ACCTGGGAGACCTGGTGGACCGACTACTGCGAGGCCACCTGGATCCCGGAGTGGGAGTTC 1920
DB 1849 ACCTGGGAGACCTGGTGGACCGACTACTGCGAGGCCACCTGGATCCCGGAGTGGGAGTTC 1908
QY 1921 GTGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGAGCCCATCATCGGC 1980
DB 1909 GTGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGAGCCCATCATCGGC 1968
QY 1981 GCGAGACCTTCTACGTGGAGCGGCGGCCAACCGCGAGACCAAGATCGGCAAGGCGCGC 2040
DB 1969 GCGAGACCTTCTACGTGGAGCGGCGGCCAACCGCGAGACCAAGATCGGCAAGGCGCGC 2028
QY 2041 TAGCTGACCCAGCCGCGCGCAGAGATCGTGAGCTGACCGAGACCAACCAAGGAG 2100
DB 2029 TAGCTGACCCAGCCGCGCGCGCAGAGATCGTGAGCTGACCGAGACCAACCAAGGAG 2088
QY 2101 ACCGAGCTGACGCGCATTCACGTGGCCCTGACGACAGCGGCGAGAGTGAACATCGTG 2160
DB 2089 ACCGAGCTGACGCGCATTCACGTGGCCCTGACGACAGCGGCGAGAGTGAACATCGTG 2148
QY 2161 ACCGACAGCAGTACGCCCTGGGCGATCATCAGGCCCAGCCCGACAGAGGAGGAGGAG 2220
DB 2149 ACCGACAGCAGTACGCCCTGGGCGATCATCAGGCCCAGCCCGACAGAGGAGGAGGAG 2208
QY 2221 CTGCTGAACAGATCATCGAGCAGCTCATCAAGAAGAGAGAGTGTACCTGAGCTGGGTG 2280
DB 2209 CTGCTGAACAGATCATCGAGCAGCTCATCAAGAAGAGAGAGTGTACCTGAGCTGGGTG 2268
QY 2281 CCGCGCCACAAAGGCGATCGCGCGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGCGATC 2340
DB 2269 CCGCGCCACAAAGGCGATCGCGCGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGCGATC 2328
QY 2341 CGCAAGTGTCTTCTTGGAGCGGATCGATGGGCGCATCGTGATCTACAGTACATGAGC 2400
DB 2329 CGCAAGTGTCTTCTTGGAGCGGATCGATGGGCGCATCGTGATCTACAGTACATGAGC 2388
QY 2401 GACCTGTACCTGGGCGCGCGCTAGGATCGATTAAAGCTTCCCGGGCTAGCAC 2460
DB 2389 GACCTGTACCTGGGCGCGCGCTAGGATCGATTAAAGCTTCCCGGGCTAGCAC 2448
QY 2461 GGTGAATTC 2469
DB 2449 GGTGAATTC 2457

RESULT 4
AX427930
LOCUS
DEFINITION
Sequence 168 from Patent WO0232943.
ACCESSION
AX427930
VERSION
AX427930.1 GI:21538017
KEYWORDS
synthetic construct.

9166 bp DNA linear PAT 20-JUN-2002

Db	3740	TCTGCACCGAGATGGGAAGAGGGCAAGATCAGCAAGATCGGCCCGAGAACCCCTACA	3799
QY	842	ACACCCCGTTTCGCCCATCAAGAAGAGCAGCACCAAGTGGCCCAAGCTGTGTGACT	901
Db	3800	ACACCCCGTTTCGCCCATCAAGAAGAGAGCAGCACCAAGTGGCGCAAGCTGTGTGACT	3859
QY	902	TCCGCGAGCTGAACAAGCGCACCAAGGACTTCTGGAGGTGTCAGCTGGGCATCCCCACC	961
Db	3860	TCCGCGAGCTGAACAAGCGCACCAAGGACTTCTGGAGGTGTCAGCTGGGCATCCCCACC	3919
QY	962	CCGCGGCCCTGAAGAAGAAAGAGCGTGACCGTCTGGAGTGGCGCAGCCTACTTCA	1021
Db	3920	CCGCGGCCCTGAAGCAGAAAGAGCGTGACCGTCTGGAGTGGCGCAGCCTACTTCA	3979
QY	1022	CGTGGCCCTGGAGAGGACTTCCGCAAGTACACCGCTTCACATCCCCAGCATCAACA	1081
Db	3980	CGTGGCCCTGGAGAGGACTTCCGCAAGTACACCGCTTCACATCCCCAGCATCAACA	4039
QY	1082	ACGAGACCCCGGCATCCGCTACCACTACCAAGTCTGCCCAAGGCTGGAAGGCGACCC	1141
Db	4040	ACGAGACCCCGGCATCCGCTACCACTACCAAGTCTGCCCAAGGCTGGAAGGCGACCC	4099
QY	1142	CCAGCATTTCCAGACGACATGACCAAGATCCTGGAGCCCTCCGGCCCGCAACCCCG	1201
Db	4100	CCGCATTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTCCGCAAGCAGAACCCCG	4159
QY	1202	AGATCGTGATCTACCACTACATGGAGACCTTACGTGGGAGCGACTGGAGATCGGCC	1261
Db	4160	ACATCGTGATCTACCACTACATGGAGACCTTACGTGGGAGCGACTGGAGATCGGCC	4219
QY	1262	AGACCGCGCAAGATCGAGGAGCTGGCGAAGCACTGCTGGCTGGGGCTTCAACACC	1321
Db	4220	AGACCGCACAGATCGAGGAGCTGGCGAAGCACTGCTGGCTGGGGCTTCAACACC	4279
QY	1322	CCGACAAGAGCACCAAGAAGAGCCCGCTTCTGTGATGGGTACGAGTGCACCCCG	1381
Db	4280	CCGACAAGAGCACCAAGAAGAGCCCGCTTCTGTGATGGGTACGAGTGCACCCCG	4339
QY	1382	ACAAGTGGACCTGCAGCCCATCGAGCTGCCGAGAAGGAGAGCTGGACGTGAACGACA	1441
Db	4340	ACAAGTGGACCTGCAGCCCATCGTCTGCCGAGAAGGAGAGCTGGACGTGAACGACA	4399
QY	1442	TCCAGAAGCTGGTGGCAAGCTGAATGGCGACCGACATCTACCCGGGATCAAGGTGC	1501
Db	4400	TCCAGAAGCTGGTGGCAAGCTGAATGGCGACCGACATCTACCCGGGATCAAGGTGC	4459
QY	1502	GCCAGCTGTCAAGCTGCTCGCGCGCGCAAGGCCCTGACCGACATCGTGCCTCTGACCG	1561
Db	4460	GCCAGCTGTCAAGCTGCTCGCGCGCACCAAGGCCCTGACCGAGTGTGCCCTGACCG	4519
QY	1562	AGAGCGCGAGCTGGAGCTGGCCGAGAACCGGGAGATCTCTCGCGACCGCTGCAGGGG	1621
Db	4520	AGAGCGCGAGCTGGAGCTGGCCGAGAACCGGGAGATCTCTGAGGAGCCGTGCACGGG	4579
QY	1622	TGTACTACGACCCAGCAGGACTTGGTGGCCGAGATCCAGAACCGGGCCACGACCACT	1681
Db	4580	TGTACTACGACCCAGCAGGACTTGGTGGCCGAGATCCAGAACCGGGCCACGACCACT	4639
QY	1682	GGACCTACCAGATCTACCAAGAGCCCTTCAAGACCTTGAAGACCGGCAAGTAGCGCAAGA	1741
Db	4640	GGACCTACCAGATCTACCAAGAGCCCTTCAAGACCTTGAAGACCGGCAAGTAGCGCGCA	4699
QY	1742	TGCGCACCGCCACCAACAGACTGAAGCAGCTGACCGAGGCGGTGCAGAACTCGCCA	1801
Db	4700	TGAAGGGCGCCACCAACAGACTGAAGCAGCTGACCGAGGCGGTGCAGAACTCGCCA	4759
QY	1802	TGAGAGCATCTGTGATCTGGGGCAGACCCCAAGTTCCGCTGCCCATCCAGAGGAGA	1861
Db	4760	CCGAGAGCATCTGTGATCTGGGGCAGACCCCAAGTTCCGCTGCCCATCCAGAGGAGA	4819
QY	1862	CTTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTGC	1921

```
Db 4820 CCTGGAGGCGCTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCGGAGTGGAGTTGG 4879
QY 1922 TGAACACCCCGCCCTGGTCAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGCG 1981
Db 4880 TGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGCG 4939
QY 1982 CGAGACCTTCTACGTGGAGCGCGCGCCCAACCGCAGAGACCAAGATCGGCAAGCGCGGT 2041
Db 4940 CCGAGACCTTCTACGTGGAGCGCGCGCCCAACCGCAGAGACCAAGTGGGCAAGCGCGGT 4999
QY 2042 ACCTGACGACCGCGCGCGCAGAGAACTGTGTAGCCTGACCGAGAGACCAACCAAGA 2101
Db 5000 ACCTGACGACCGCGCGCGCAGAGAACTGTGTAGCCTGACCGAGAGACCAACCAAGA 5059
QY 2102 CGAGCTGAGGCGCATCAGCTGGCCCTGCAGGACAGCGGACCGAGGTGAACATCGTGA 2161
Db 5060 CGAGCTGAGGCGCATCAGCTGGCCCTGCAGGACAGCGGCTGGAGGTGAACATCGTGA 5119
QY 2162 CGACAGCCAGTACGCGCTGGGCGATCATCCAGGCCAGCGCCGCAAGAGCGAGAGCGAGC 2221
Db 5120 CGACAGCCAGTACGCGCTGGGCGATCATCCAGGCCAGCGCCGCAAGAGCGAGAGCGAGC 5179
QY 2222 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGAGAGGTGTACCTGAGCTGGGTGC 2281
Db 5180 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGAGAGGTGTACCTGAGCTGGGTGC 5239
QY 2282 CGCCCGCAGGCGCATCGCGCGCAAGCAGCAGATCGACAAGCTGGTCAGCAAGGCGATCC 2341
Db 5240 CGCCCGCAGGCGCATCGCGCGCAAGCAGCAGATCGACAAGCTGGTCAGCAAGGCGATCC 5299
QY 2342 GCAAGTGTCTTCTCGAGCGGCGATCGA 2369
Db 5300 GCAAGTGTCTTCTCGAGCGGCGATCGA 5327

RESULT 5
AX427936
LOCUS AX427936 9788 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 174 from Patent WO0232943.
ACCESSION AX427936
VERSION AX427936.1 GI:21538023
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 174 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Source
1. .9788
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2377 a 2817 c 2695 g 1899 t
ORIGIN
Query Match 82.7%; Score 2042.8; DB 6; Length 9788;
Best Local Similarity 91.8%; Pred. No. 5e-206;
Matches 2197; Conservative 0; Mismatches 177; Indels 18; Gaps 3;

QY 14 TGGCCGAGGCGCATGAGCCAGCGCCAGCCCAACATCTCTGATGCGAGCGCAACTTCA 73
Db 2961 TGGCCGAGGCGCATGAGCAAGGTGAACCAACCAACATCATGATGCGAGCGCAACTTCA 3020
QY 74 AGGCCCGCAAGCGCATCATCAAGTGTCTCAACTCGGCAAGGAGGGCCACATCGCCGCA 133
Db 3021 AGGCCCGCAAGCGCATCATCAAGTGTCTCAACTCGGCAAGGAGGGCCACCTGGCCGCA 3080
QY 134 ACTGCGCGCGCCCGCAAGAGGCTGTGGAAGTGCAGCAAGGAGGCGCCACAGATGA 193
```

```
Db 3081 ACTGCGCGCGCCCGCCCGCAAGAGAGGGCTGCTGGAAGTGCAGCAAGAGGCGCCACAGGTGA 3140
QY 194 AGGACTGCAACGAGCGCCAGCGCCCAACTTCTTCGCGGAGGACCTGGCTTCCCCAGGGCA 253
Db 3141 AGGACTGCAACGAGCGCACAGGCTAA-----TAGGGAAGATCTGGGCTTCCCCAAGGGA 3194
QY 254 AGCCCGCGGAGTTTCCCGCAGCAGAGCAACCGCCCAACAGCCCGCCAGCGCGGAGGTGC 313
Db 3195 AGCCCGGAGGAATTTCTTCAGAGCAGACAGAGCCCAACAGCCCGCCAGAGAGAGCTTC 3254
QY 314 AGTGCGCGG-----CGACAACCCCGCAGAGGCGCGCGGCGAGCGCGCAGGGCA--- 364
Db 3255 AGTGTGGGGAAGAGACAACAACCTCCCTCTCAGAAGCAGGAGCGGATAGACAAGAACTG 3314
QY 365 ---CCCTGAACTTCCCGCCAGATCACCTGTGGCAGCGCCCTTGGTGAGCATCAAGTGG 421
Db 3315 TATCTTTTACCTTCCCTCAGATCCTTTTGGCAGCAGCCCTCGTGCACATAAAGATAG 3374
QY 422 GCGGCCAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGCAGACACCGTGTCTGGAGGGA 481
Db 3375 GGGGCCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGCAGACACCGTGTCTGGAGGGA 3434
QY 482 TGAGCCTGCCCGGGAAGTGAAGCCCAAGATGATCGCGCGGCGATCGCGGCTTCAATCAGG 541
Db 3435 TGAACCTGCCCGCGCGCTGGGAAGCCCAAGATGATCGCGCGGCGATCGCGGCTTCAATCAGG 3494
QY 542 TGCGCCAGTACGACACCATCTGATCGAGATCTCGGCAAGAGGCGCATCGGACCGTGC 601
Db 3495 TGGCCAGTACGACACCATCTGATCGAGATCTGCGGCCACAGGCGCATCGGACCGTGC 3554
QY 602 TGATCGCCCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCCAGTGGGCTGCA 661
Db 3555 TGGTGGSCCGCCACCGCGTGAACATCATCGCGCGCAACCTGCTGACCCAGATCGGCTGCA 3614
QY 662 CCCTGAACCTTCCCGATCAGCCCGATCGAGACCGTGCCTGCGGCGGATGG 721
Db 3615 CCCTGAACCTTCCCGATCAGCCCGATCGAGACCGTGCCTGCGGCGGATGG 3674
QY 722 ACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCGCTGACCGCCA 781
Db 3675 ACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCGCTGCTGGAGA 3734
QY 782 TCTCGGAGGAGTGGAGAAGGCGGCAAGATCAACAGATCGGCGCGCGAGAACCCCTACA 841
Db 3735 TCTGCAACGAGATGGAGAAGGCGGCAAGATCAGCAAGATCGGCGCGCGAGAACCCCTACA 3794
QY 842 ACACCCCGCTGTTCGCCATCAAGNAGAGACACACCAAGTGGCGCAAGTGTGACT 901
Db 3795 ACACCCCGCTGTTCGCCATCAAGNAGAGAGACACCAAGTGGCGCAAGTGTGACT 3854
QY 902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGCTGGGATCCCCCACC 961
Db 3855 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGCTGGGATCCCCCACC 3914
QY 962 CCGCGCGCTTGAAGAAGAAGAGCGTGCCTGCTGGAGTGGCGGCGACCGCTACTTCA 1021
Db 3915 CCGCGCGCTTGAAGAAGAAGAGCGTGCCTGCTGGAGTGGCGGCGACCGCTACTTCA 3974
QY 1022 GCGTGCCTTGGAGGAGGACTTCCGCAAGTACACCGCGCTTCCACCATCCCGAGCATCAACA 1081
Db 3975 GCGTGCCTTGGAGGAGGACTTCCGCAAGTACACCGCGCTTCCACCATCCCGAGCATCAACA 4034
QY 1082 ACAGAGCCCCCGGCTCCGCTACCCAGTACACGCTGCTGCGCGCGGCTGGAGGCGAGCC 1141
Db 4035 ACAGAGCCCCCGGCTCCGCTACCCAGTACACGCTGCTGCGCGCGGCTGGAGGCGAGCC 4094
QY 1142 CCAGCATTTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCCGCGCGCGCGCGCG 1201
Db 4095 CCGCATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCGCTTCCCGAAGCAGAACCCCG 4154
QY 1202 AGATCGTGTATCTACCATGATGACACCGTGTAGTGGGAGGCGAGCTGGAGATCGGCG 1261
Db 4155 ACATCGTGTATCTACCATGATGACACCGTGTAGTGGGAGGCGAGCTGGAGATCGGCG 4214
```


Best Local Similarity 91.6%; Pred. No. 2.3e-204; Matches 2194; Conservative 0; Mismatches 185; Indels 16; Gaps 4;			
QY	14	TGCGCGAGGCGCATGAGCCAGG---CCACGAGCGCCCAACATCTGTGATGACGCGAGCAACT	70
Db	2967	TGCGCGAGGCGCATGAGCCAGGTCACCAAGCGCCACCACATCATGATGACGCGGGAACCT	3026
QY	71	TCAGAGGCGCCCAAGCGCATCATCAAGTGTTCAACTGCGGCAAGGAGGCGCCACATCGCCC	130
Db	3027	TCGCAACCCAGCGCAAGATCGTGAAGTGTTCAACTGCGCAAGGAGGCGCCACACCGCCC	3086
QY	131	GCNACTGCGCGCGCCCGCCGCAAGAGGGTGTGTGAAGTGTGCGCAAGGAGGCGCCACCA	190
Db	3087	GCAACTGCGCGCGCCCGCCGCAAGAGGGTGTGTGAAGTGTGCGCAAGGAGGCGCCACCA	3146
QY	191	TGAAGGACTGCAACGAGCGCCAGGCGCAACTTCTCCGCGAGGAGGAGGAGGCGCCAGG	250
Db	3147	TGAGGAGCTGCAACGAGCGCCAGGCGCAACTTCTCCGCGAGGAGGAGGAGGCGCCAGG	3205
QY	251	GCAAGCGCGCGAGTTCGCCAGCGAGCAGACGCGCCCAACAGCGCCACAGCGCGGAGC	310
Db	3206	GGAAGCGCCAGGGAATTTCTTACAGAGCAGACAGAGCCCAACAGCGCCACAGAGAGAGC	3265
QY	311	TGCAGTGTGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGCGCGAGCGGCA	364
Db	3266	TTCAGGTTTGGGGAAGAGACAACACTCCCTCTCAGAAAGCAGGAGCGCGATGACAAAGAA	3325
QY	365	-----CCCTGAACCTTCCCGCCAGATCACCTGTGGCAGCGCCCTGTGGTGAAGCATCAAG	418
Db	3326	CTGTATCTTTAGCTTCCCTCAGATCAGTCTTTGGCAGCGACCCCTCTGCAATAAAGA	3385
QY	419	TGGCGCGCGAGATCAAGGAGGCGCTGTGGACACCGCGCGCAGACACCGTGTGGAGG	478
Db	3386	TAGGGCGCGCAGTGAAGGAGGCGCTTCTAGACACCGCGCGCAGACCGTGTGGAGG	3445
QY	479	AGATGAGCTGTCCCGCAAGTGAAGCCCAAGATGATCGGGCGCATCGGGCGCTTATCA	538
Db	3446	AGATGAACCTGTCCCGCGCGCTGGAAGCCCAAGATGATCGGGCGCATCGGGCGCTTATCA	3505
QY	539	AGGTGCGCGCAGTACGACCAAGTCTTATCGAGATCTGCGGCAAGAGGCGCATCGGCACCG	598
Db	3506	AGGTGCGCGCAGTACGACCAAGTCTTATCGAGATCTGCGGCAAGAGGCGCATCGGCACCG	3565
QY	599	TGCTGATCGGCGCGCGCGCGTGAACATCATCGGCGCAACATGTGACCCAGCTGGGT	658
Db	3566	TGCTGATCGGCGCGCGCGCGTGAACATCATCGGCGCAACATGTGACCCAGCTGGGT	3625
QY	659	GCACCTGAACCTTCCCGCATCGCCCATCGAGACCGTGCCTGTGAAGCTGAAGCCCGGCA	718
Db	3626	GCACCTGAACCTTCCCGCATCGCCCATCGAGACCGTGCCTGTGAAGCTGAAGCCCGGCA	3685
QY	719	TGAGCGCGCCCAAGGTGAAGCAGTGGCGCGCTGACCGAGGAGAGATCAAGGCGCTGACCG	778
Db	3686	TGAGCGCGCCCAAGGTGAAGCAGTGGCGCGCTGACCGAGGAGAGATCAAGGCGCTGAGG	3745
QY	779	CCATCTGCGAGGAGATGGAAGAGGAGGCAAGATCACCAAGATCGGCGCCCGAGAACCCCT	838
Db	3746	AGATCTGCAACCGAGATGGAAGAGGAGGCAAGATCAGCAAGATCGGCGCCCGAGAACCCCT	3805
QY	839	ACAACACCCCGCTGTTCGCCATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	898
Db	3806	ACAACACCCCGCTGTTCGCCATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	3865
QY	899	ACTTCCGCGAGTGAACAGCGCACCCAGGAGCTTCTGGGAGGTGAGCTGGGCATCCCCC	958
Db	3866	ACTTCCGCGAGTGAACAGCGCACCCAGGAGCTTCTGGGAGGTGAGCTGGGCATCCCCC	3925
QY	959	ACCCCGCGCGCTGAGAGAGAGAGAGCGGTGACCGTGTGACGCTGGGCGGACGCGCTACT	1018
Db	3926	ACCCCGCGCGCTGAGAGAGAGAGAGCGGTGACCGTGTGACGCTGGGCGGACGCGCTACT	3985
QY	1019	TCAGGCGTCCCGTGGACGAGGAGTTCGCAAGTACACCGCTTCACCATCCCGCAGCATCA	1078
Db	5066	AGACCGAGCTGCAGGCGCATCCACCTGGCGCTGCAGGACAGCGGCGCTGGAGGTGAACATCG	5125
Db	3986	TCAGCGTGGCCCTGGAGCAAGGACTTCCGCAAGTACACCGCTTACCATCCCGCAGCATCA	4045
QY	1079	ACAACGAGACCCCGCGCATCCGCTACAGTACAACTGTGTCGCCAGGCTGGAAGGGCA	1138
Db	4046	ACAACGAGACCCCGCGCATCCGCTACAGTACAACTGTGTCGCCAGGCTGGAAGGGCA	4105
QY	1139	GCCCCAGCATTTCCAGAGCAGCATGACCAAGATCTTGGAGCCCTTCCCGCGCCCGCAACC	1198
Db	4106	GCCCCAGCATTTCCAGTGCAGCATGACCAAGATCTTGGAGCCCTTCCCGCAAGCAAGACC	4165
QY	1199	CGAGATCTGTGATTTACCACTACATGACGACCTGTACGTGGGCGAGGAGCTGGAGATCG	1258
Db	4166	CGACATCTGTGATTTACCACTACATGACGACCTGTACGTGGGCGAGGAGCTGGAGATCG	4225
QY	1259	GCAGACACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGGCTTCAACA	1318
Db	4226	GCAGACACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGGCTTCAACA	4285
QY	1319	CCCCGACAAAGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1378
Db	4286	CCCCGACAAAGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	4345
QY	1379	CGACAAGTGGAGCTGACCGCATCGAGTGCCTCGAGAGGAGGAGGAGGAGGAGGAGGAGG	1438
Db	4346	CGACAAGTGGAGCTGACCGCATCGAGTGCCTCGAGAGGAGGAGGAGGAGGAGGAGGAGG	4405
QY	1439	ACATCCAGAGCTGTGGCAAGCTGAACCTGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGG	1498
Db	4406	ACATCCAGAGCTGTGGCAAGCTGAACCTGGGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGG	4465
QY	1499	TGCGCGCAGCTGTCAAGCTGCTGCGCGGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1558
Db	4466	TGCGCGCAGCTGTCAAGCTGCTGCGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	4525
QY	1559	CGAGAGGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGAGTCTTGGCGGAGGAGGAGGAGGAGG	1618
Db	4526	CGAGAGGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGAGTCTTGAAGGAGGAGGAGGAGGAGG	4585
QY	1619	GGGTGTACTACGACCCAGCAAGGAGCTGTGGCGCGAGATCCAGAGAGGAGGAGGAGGAGGAGG	1678
Db	4586	GGGTGTACTACGACCCAGCAAGGAGCTGTGGCGCGAGATCCAGAGAGGAGGAGGAGGAGGAGG	4645
QY	1679	AGTGGACCTTACAGATCTTACAGGAGGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA	1738
Db	4646	AGTGGACCTTACAGATCTTACAGGAGGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA	4705
QY	1739	AGATGCGCACCGCCCGCACCAAGGAGCTGAAGCAGCTGACCGAGGCGCTGACAGAAATCG	1798
Db	4706	GCATGAAGGCGCGCCCGCACCAAGGAGCTGAAGCAGCTGACCGAGGCGCTGACAGAAATCG	4765
QY	1799	CCATGGAGAGCATCGTGTGTTGGGCAAGACCCCGCAAGTTCCGCTTCCCATCCAGAAAG	1858
Db	4766	CCACCGAGAGCATCGTGTGTTGGGCAAGACCCCGCAAGTTTCAAGCTGCCCATCCAGAAAG	4825
QY	1859	AGACCTGGGAGACCTGGTGGACCGACTACTGGAGCGCCACCTGGATGCCCGAGTGGGAGT	1918
Db	4826	AGACCTGGGAGGCGTGGTGGACCGAGTACTGGAGCGCCACCTGGATGCCCGAGTGGGAGT	4885
QY	1919	TGCTGAACACCCCGCTGGTGAAGCTGTGTACAGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGG	1978
Db	4886	TCGTGAACACCCCGCTGGTGAAGCTGTGTACAGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGG	4945
QY	1979	GGCGCGAGACCTTCTACGTGGAGCGCGCGCAACCGCGAGACCAAGATCGGCAAGGCGG	2038
Db	4946	GGCGCGAGACCTTCTACGTGGAGCGCGCGCAACCGCGAGACCAAGATCGGCAAGGCGG	5005
QY	2039	GCTAGCTGACCGAGCG	2098
Db	5006	GCTAGCTGACCGAGCG	5065
QY	2099	AGACCGAGCTGCAGGCGCATCCAGCTGGCGCTGCAGGACAGCGCGCGCGCGCGCGCGCGCG	2158
Db	5066	AGACCGAGCTGCAGGCGCATCCACCTGGCGCTGCAGGACAGCGCGCGCGCGCGCGCGCGCG	5125

QY 1499 TGGCCAGCTGTGCAAGCTGCTGCGGGCGCCCAAGGCCCTGACCGACATCGTCCCTGTA 1558
Db 4466 TGGCCAGCTGTGCAAGCTGCTGCGGGCGCCCAAGGCCCTGACCGAGGTGGTCCCTGTA 4525
QY 1559 CCAGAGGCGGAGCTGGAGCTGGCGAGAACCGCGAGATCCCTGCGCGAGCCGTCGACG 1618
Db 4526 CCAGAGGCGGAGCTGGAGCTGGCGAGAACCGCGAGATCCCTGGAAGGAGCCGTCGACG 4585
QY 1619 GCGTGTACTACGACCCACGACGAGCTGTGGCGGAGATCCAGAGATCCAGAGCGGCGACGACC 1678
Db 4586 GCGTGTACTACGACCCACGACGAGCTGTGGCGGAGATCCAGAGATCCAGAGCGGCGACG 4645
QY 1679 AGTGGACCTACGAGATCTACGAGAGGCGCTTCAAGAACCTGGAAGACCGGCAATACGCCA 1738
Db 4646 AGTGGACCTACGAGATCTACGAGAGGCGCTTCAAGAACCTGGAAGACCGGCAATACGCC 4705
QY 1739 AGATGCGCACCGCCACGACGAGCTGTGAAGCAGCTGACCGAGGCGGTGCAGAAATCG 1798
Db 4706 GATGAAGGCGGCCACCAACGAGCTGAAGCAGCTGACCGAGGCGGTGCAGAAATCG 4765
QY 1799 CCATGGAGAGCATCGTGTGCTGGGCAAGACCCCAAGTTCCCGCTGCCCATCCAGAAGG 1858
Db 4766 CCACGAGAGCATCGTGTGCTGGGCAAGACCCCAAGTTCAAGCTGCCATCCAGAAGG 4825
QY 1859 AGACCTGGGAGACCTGGTGGACGAGTACTTGGCAGGCGACCTGGATCCCGAGTGGAGT 1918
Db 4826 AGACCTGGGAGGCGTGGTGGACGAGTACTTGGCAGGCGACCTGGATCCCGAGTGGAGT 4885
QY 1919 TCCTGAACACCCCGCTGGTGAAGCTGTGTAACGCTGGAGAGGAGGCCCATCATCG 1978
Db 4886 TCCTGAACACCCCGCTGGTGAAGCTGTGTAACGCTGGAGAGGAGGCCCATCATCG 4945
QY 1979 GCGCCGAGACCTTCTACGTGGAGGCGCGGCCCAACCGCGAGACCAAGATCGCGAAGGCCG 2038
Db 4946 GCGCCGAGACCTTCTACGTGGAGGCGCGGCCCAACCGCGAGACCAAGTGGGCAAGGCCG 5005
QY 2039 GCTACGTGACCGAGCGGCGCGCAAGATGCTGAGCCTGACCGAGACCAACCAACAGA 2098
Db 5006 GCTACGTGACCGAGCGGCGCGCAAGATGCTGAGCCTGACCGAGACCAACCAACAGA 5065
QY 2099 AGACCGAGCTGACGAGCATCCAGCTGGCCCTGCAGGACACGCGGCGAGGCTGAACATCG 2158
Db 5066 AGACCGAGCTGACGAGCATCCAGCTGGCCCTGCAGGACACGCGGCTGGAGTGAACATCG 5125
QY 2159 TGACCGACAGCGAGTACGCGCTGGGCATCATCCAGCGCCGAGCCGACACAGAGCGAGCGG 2218
Db 5126 TGACCGACAGCGAGTACGCGCTGGGCATCATCCAGCGCCGAGCCGACACAGAGCGAGCG 5185
QY 2219 AGCTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGTGTACCTGAGCTGGG 2278
Db 5186 AGCTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGTGTACCTGAGCTGGG 5245
QY 2279 TGCCCGCCCAAGGCGATCGCGCGCAAGCAGAGATCGACAGCTGGTGAAGGCGCA 2338
Db 5246 TGCCCGCCCAAGGCGATCGCGCGCAAGCAGAGTGTGACGCGCTGGTGAAGGCGCGCA 5305
QY 2339 TCGGCAAGTGTGTTCTTGGAGCGGATCGATGGCGGCGATCGTGTACCAAGTA 2393
Db 5306 TCGGCAAGTGTGTTCTTGGAGCGGATCGACAAAGGCCCGAGGAGGAGCAGAGAA 5360

RESULT 9
AX427926
LOCUS AX427926 9194 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 164 from Patent WO0232943.
ACCESSION AX427926
VERSION AX427926.1 GI:21538013
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE artificial sequences.
1
AUTHORS Huang, Y. and Nabel, G.J.

TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 164 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES Location/Qualifiers
source 1. .9194
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2232 a 2672 c 2523 g 1767 t
ORIGIN
Query Match 82.0%; Score 2025.4; DB 6; Length 9194;
Best Local Similarity 91.6%; Pred. No. 3.4e-204;
Matches 2193; Conservative 0; Mismatches 186; Indels 16; Gaps 4;
QY 14 TGGCCAGGCGCATGAGCCAGG---CCACAGCGCCCAACATCTCTGTATGTCAGCGGCAACT 70
Db 2967 TGGCCAGGCGCATGAGCCAGGTCACCAACAGCGCCACATCATCATGTCAGCGGCGCAACT 3026
QY 71 TCAAGGGCGCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGCGCCACATCGCCC 130
Db 3027 TCCGCAACACCGCCAGATCTGTAAGTGTCTCAACTGCGGCAAGGAGGCGCCACCGCCC 3086
QY 131 GCAACTGCCGCGCCCGCCCGCAAGAAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACACAGA 190
Db 3087 GCAACTGCCGCGCCCGCCCGCAAGAAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACACAGA 3146
QY 191 TGAAGACTGTCACCGAGCGCCAGCCAACTTCTTCGCGAGGACCTGGCCCTTCCCCCAGG 250
Db 3147 TGAAGACTGTCACCGAGCGACAGGCTAA-TTTTATTAGGAAGATCTGGCCCTTCCCCAAG 3205
QY 251 GCAAGCGCGCGAGTTTCCCGAGCGAGCAACCGCGCAACACGCCCGCCACCGCGCGAGC 310
Db 3206 GGAAGCGCGAGGAAATTTCTTCAGAGCAGACAGACAGCCACAGCCCGCCACCAAGAGAGC 3265
QY 311 TGCAGTGGCGCG- - - - -CGACAACCCCGCGAGCGAGGCGCGCGCGAGCGCGCGAGGCA 364
Db 3266 TTCAGGTTGGGGAAGAGACAACACTCCCTCTCAGAAAGCAGGAGCGGATAGACAAGGAA 3325
QY 365 - - - - -CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGGTGAAGTCAAGG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTGCACAAATAAGA 3385
QY 419 TGGCGCGCGCAGATCAAGGAGCGCCCTGCTGGACACCGCGCGCGAGCACACCGTCTGGAGG 478
Db 3386 TAGGCGCGCAGCTGAAGAGGCGCCCTTCTAGACACCGCGCGCGAGCACACCGTCTGGAGG 3445
QY 479 AGATGAGCTGCCCGCAAGTGGAAAGCCCAAGATGATCGGCGGCATCGGGCGGCTTCATCA 538
Db 3446 AGATGAACCTGCGCGCGCGTGGAAAGCCCAAGATGATCGGCGGCATCGGGCGGCTTCATCA 3505
QY 539 AGGTGCGCGAGTACGACCGAGATCTCTGATCGAGATCTTGGCGCAAGAGGCCATCGGCACCG 598
Db 3506 AGGTGCGCGAGTACGACCGAGATCTCTGATCGAGATCTTGGCGCCACAAGGCCATCGGCACCG 3565
QY 599 TGTGTATCGCGCCCGCCCGCTGAACATCATCGGCGCGCAACATCTGACCCAGCTTGGGCT 658
Db 3566 TGTGTGTGGCGCCCGCCCGCTGAACATCATCGGCGCGCAACCTGCTGACCCAGATCGGCT 3625
QY 659 GCACCTGAACTTCCCATCAGCCCGCATCGAGACCTGCGCGGTGAAGCTGAAGCGCGGCA 718
Db 3626 GCACCTGAACTTCCCATCAGCCCGCATCGAGACCTGCGCGGTGAAGCTGAAGCGCGGCA 3685
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCCAGGAGAGATCAAGGCCCTGACCG 778
Db 3686 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCCAGGAGAGATCAAGGCCCTGCTGGT 3745
QY 779 CCATCTGCGAGGAGATGGAGAAGGCGCAAGATCAACCAAGATCGGCCCGCGGAGACCCCT 838
Db 3746 AGATCTGCGACCGAGATGGAGAAGGCGCAAGATCAACCAAGATCGGCCCGCGGAGACCCCT 3805
QY 839 ACAACACCCCGCTGTTCCGCATCAAGAAAGGACAGACCAAGTGGCGCAAGCTGGTG 898

|||||
Db 3806 ACACACCCCGTTTCGCCATCAAGAGAGACAGCACCAAGTGGCGAAGCTGGTG 3865
QY 899 ACTTCGCGAGCTGACAAAGCGCACCCAGGACTTCTGGGAGGTGACAGTGGGATCCCC 958
|||||
Db 3866 ACTTCGCGAGCTGACAAAGCGCACCCAGGACTTCTGGGAGGTGACAGTGGGATCCCC 3925
QY 959 ACCCGCGCGCTGAAGAAAGAGCGTGACCGTGTGGAGCTGGGGGACGCGCTACT 1018
|||||
Db 3926 ACCCGCGCGCTGAAGAAAGAGCGTGACCGTGTGGAGCTGGGGGACGCGCTACT 3985
QY 1019 TCAGCGTGGCCCTGGAGAGGACTTCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 1078
|||||
Db 3986 TCAGCGTGGCCCTGGAGAGGACTTCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 4045
QY 1079 ACAACAGAGACCCCGGATCCGCTACAGTACAAAGTGTGCGCCAGGGGTGGAAGGCA 1138
|||||
Db 4046 ACAACAGAGACCCCGGATCCGCTACAGTACAAAGTGTGCGCCAGGGGTGGAAGGCA 4105
QY 1139 GCGCCAGCATCTCCAGAGCAGATGACCAAGATCTCTGGAGCGCTTCCGCGCCCGCAACC 1198
|||||
Db 4106 GCGCCGCGATCTCCAGTGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCAAGCAGAAC 4165
QY 1199 CGAGATTCGTATACCAAGTACATGACGACCTGTACGTGGGAGGAGCTGAGATCG 1258
|||||
Db 4166 CGACATCGTGATACCAAGTACATGACGACCTGTACGTGGGAGGAGCTGAGATCG 4225
QY 1259 GCAGAGACCCCGGATCGAGAGCTGCGAAGCAGCTGTGCGGTGGGGCTTACCA 1318
|||||
Db 4226 GCCAGCACCCGACAGATCGAGAGCTGCGCAGCAGCTGTGCGGTGGGGCTTACCA 4285
QY 1319 CCGCCGACAGAGACACCAAGAGCGCCCTTCTCTGTGGATGGGCTACGAGCTGCACC 1378
|||||
Db 4286 CCGCCGACAGAGACACCAAGAGCGCCCTTCTCTGTGGATGGGCTACGAGCTGCACC 4345
QY 1379 CGACAAAGTGGACCGTGCAGCCCATCGAGCTGCGCGAGAGGAGCTGGAACG 1438
|||||
Db 4346 CGACAAAGTGGACCGTGCAGCCCATCGAGCTGCGCGAGAGGAGCTGGAACG 4405
QY 1439 ACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCGGCATCAAG 1498
|||||
Db 4406 ACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCGGCATCAAG 4465
QY 1499 TGGCGCAGCTGTCAAGCTGCTGCGGGCGCAAGGCGCTGACGACATCGTGCCTGTA 1558
|||||
Db 4466 TGGCGCAGCTGTCAAGCTGCTGCGGGCGCAAGGCGCTGACGAGGTGGTGCCTGTA 4525
QY 1559 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGCGGAGCCCGTGCACG 1618
|||||
Db 4526 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGAGGAGCCCGTGCACG 4585
QY 1619 GCGTGTACTACGACCCAGCAAGGACCTGTGGCCGAGATCCAGAACGCGGCGCACACC 1678
|||||
Db 4586 GCGTGTACTACGACCCAGCAAGGACCTGTGGCCGAGATCCAGAACGCGGCGGAGGCC 4645
QY 1679 AGTGGACCTACCATCTACAGAGCGCTTCAAGAACCTGAAGACCGCAAGTACGCCA 1738
|||||
Db 4646 AGTGGACCTACCATCTACAGAGCGCTTCAAGAACCTGAAGACCGCAAGTACGCCA 4705
QY 1739 AGATGCGCACCGGCCACACCAACGACGTGAAGCAGCTGACCGAGGCGCTGCAGAAATCG 1798
|||||
Db 4706 GCATGAAGGGCGGCCACACCAACGACGTGAAGCAGCTGACCGAGGCGCTGCAGAAATCG 4765
QY 1799 CCATGGAGAGCATCGTGATCTGGGGCAAGACCCCGCAAGTTCGCGCTCGCCATCCAGAGG 1858
|||||
Db 4766 CCACCTGGAGAGCATCGTGATCTGGGGCAAGACCCCGCAAGTTCGAGCTCGCCATCCAGAGG 4825
QY 1859 AGACCTGGGAGACCTGGTGACCGACTACTGGCAGGCGCACTGGATCCCGAGTGGGAGT 1918
|||||
Db 4826 AGACCTGGGAGGCTGGTGACCGAGTACTGGCAGGCGCACTGGATCCCGAGTGGGAGT 4885
QY 1919 TCGTGAACACCCCGCTGGTGAAGCTGTGGTACCAAGCTGGGAAGGAGGCCCATCATCG 1978
|||||

Db 4886 TCGTGAACACCCCGCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGGCCATCATCG 4945
QY 1979 GCGCGAGACCTTCTAGCTGGACGGCGCCGCAACCCCGAGAGACCAAGATGGCAAGGCCG 2038
|||||
Db 4946 GCGCGAGACCTTCTAGCTGGACGGCGCCGCAACCCCGAGAGACCAAGCTGGCAAGGCCG 5005
QY 2039 GCTACGTGACCGACCGGGCGCGCAGAAAGTCTGAGCCCTGACCGAGACCAACCAACAGA 2098
|||||
Db 5006 GCTACGTGACCGACCGGGCGCGCAGAAAGTCTGAGCCCTGACCGAGACCAACCAACAGA 5065
QY 2099 AGACCGAGCTGCAAGGCCATCCAGCTGCGCCCTGACGAGACGGCGAGGTGAACATCG 2158
|||||
Db 5066 AGACCGAGCTGCAAGGCCATCCACCTGGCCCTGACGAGACAGCGGCTGGAGTGAACATCG 5125
QY 2159 TGACCGACACCGAGTACGCGCTGGGCATCATCCAGGCGCCAGCCGACAAAGAGCGAGAGCG 2218
|||||
Db 5126 TGACCGACACCGAGTACGCGCTGGGCATCATCCAGGCGCCAGCCGACAAAGAGCGAGAGCG 5185
QY 2219 AGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGTGTACCTGAGCTGGG 2278
|||||
Db 5186 AGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGTGTACCTGGCCTGGG 5245
QY 2279 TGCCCGCCCAAGAGGCGATCGCGGCAACGAGCAGATCGCAAGCTGTGAGCAAGGGCA 2338
|||||
Db 5246 TGCCCGCCCAAGAGGCGATCGCGGCAACGAGCAGTGGAGCGCTGTGAGCGCGCGCA 5305
QY 2339 TCCGCAAGGTGCTGTCTCTGGACGGCATCGATGCGCGCATCGTGTATACAGTA 2393
|||||
Db 5306 TCCGCAAGGTGCTGTCTCTGGACGGCATCGACAAAGGCCAGGAGGAGCAGAGAA 5360
|||||

RESULT 10
AX427938
LOCUS AX427938 9785 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 176 from Patent WO0232943.
ACCESSION AX427938
VERSION AX427938.1 GI:21538025
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
JOURNAL genetic immunization
PATENT: WO 0232943-A 176 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
source
1..9785
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2360 a 2843 c 2731 g 1851 t
ORIGIN

Query Match 81.8%; Score 2020.4; DB 6; Length 9785;
Best Local Similarity 91.6%; Pred. No. 1.1e-203;
Matches 2190; Conservative 0; Mismatches 181; Indels 21; Gaps 4;

QY 14 TGGCCGAGGCGCATGAGCCAGGCCACCAGCCATCTCTGATGAGCGCAGCAACTTCA 73
|||||
Db 2967 TGGCCGAGGCGCATGAGCCAGGTGCAGCACACCAACATCATGATGAGCGCGCAACTTCA 3026
|||||

QY 74 AGGCGCCCAAGCGCATCATCAAGTGTCTTCACTCGGGCAAGAGGGCCACATCGCCCGCA 133
|||||
Db 3027 AGGCGCGCAAGCG---CATCAAGTGTCTTCACTCGGGCAAGAGGGCCACCTTGCCCGCA 3083
|||||

QY 134 ACTCCGCGCGCCCGCGCAAGAGGCGTGTGGAAGTGCAGCAAGGAGGGCCACCAAGATGA 193
|||||
Db 3084 ACTCCGCGCGCTCGGCAAGAGGGCTGTGGAAGTGCAGCAAGGAGGGCCACCAAGATGA 3143
|||||

QY 194 AGGACTCACCGAGCGCGCAGGCCCAACTTCTCCGCGAGGAGCTGGCCTTCCCCCGAGGCA 253
|||||

Db 3144 AGGACTGCACCGAGGACAGAGGCTAA-----TAGGGAAGATCTGGGCTTCCCCACAAGGGA 3197
QY 254 AGCGCCGCGAGTTCCCCAGCAGCAGAAACCGCGCCCAACAGCCGCCACAGCGCGAGCTGC 313
Db 3198 AGCCACAGGAATTTCTCAGAGCAGCAGCAGAGCCACACAGCCGCCACAGAGAGAGCTTC 3257
QY 314 AGGTGCGCG- ----CGACACCCCGCAGCAGAGCGCGCGCGCAGCGCCAGGGCA--- 364
Db 3258 AGGTTGGGGAAGAGACAACAATCCCTCTCAGAAGCAGGAGCGCATAGACAAGGAAGTCTG 3317
QY 365 ---CCCTGAACCTCCCGCCAGATCACCTGTGGCAGCGCCCTCGTGGAGCATCAAGGTGG 421
Db 3318 TATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAATAAAGATAG 3377
QY 422 GCGGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCCGACACACCGCTGCTGGAGGAGA 481
Db 3378 GGGCCAGCTGAAGGAGGCCCTTCTAGACACCGCGCGCGGACACCGCTCTGGAGGAGA 3437
QY 482 TGAGCCTGCGCGCAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGG 541
Db 3438 TGAACCTGCGCGCGCGCTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAG 3497
QY 542 TGGCCAGTACGACACAGATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCAACGTGC 601
Db 3498 TGGCCAGTACGACACAGATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCAACGTGC 3557
QY 602 TGATCGGCGCCCGCCCGTGAACATCATCGGCGCAACATGCTGACCCAGCTGGGTGCA 661
Db 3558 TGGTGGCGCCCGCCCGTGAACATCATCGGCGCAACCTGCTGACCCAGATCGGCTGCA 3617
QY 562 CCTGAACCTCCCGCATACGCCCATCGACACCGTGCCTGTAAGCTGAAGCGCGGATGG 721
Db 3618 CCGTGAACCTCCCGCATACGCCCATCGACACCGTGCCTGTAAGCTGAAGCGCGGATGG 3677
QY 722 ACGGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCA 781
Db 3678 ACGGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGGTGGAGA 3737
QY 782 TGTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTACA 841
Db 3738 TGTGCAACGAGATGGAGAAGGAGGCAAGATCAGCAAGATCGGCCCGCGAGAACCCCTACA 3797
QY 842 ACACCCCGCTGTGCGCATCAAGAGAAGGACAGACCAAGTGGCGCAAGCTGGTGACCT 901
Db 3798 ACACCCCGCTGTGCGCATCAAGAGAAGGAGACACCAAGTGGCGCAAGCTGGTGACCT 3857
QY 902 TCGCGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGCAAGCTGGGCACTCCCGCAC 961
Db 3858 TCGCGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGCAAGCTGGGCACTCCCGCAC 3917
QY 962 CCGCGGCTGAAGAGAAGAGAGCGTGACCGTGTGACGTGGCGCAGCGCTACTTCA 1021
Db 3918 CCGCGGCTGAAGCAGAGAAGAGCGTGACCGTGTGACGTGGCGCAGCGCTACTTCA 3977
QY 1022 GGTGCGCTGGACAGGAGCTTCGCAAGTACACCGCTTCACCATCCCGCAGCATCAACA 1081
Db 3978 GGTGCGCTGGACAGGAGCTTCGCAAGTACACCGCTTCACCATCCCGCAGCATCAACA 4037
QY 1082 ACAGAGACCCCGCGCATCCGCTACCAAGTACAAGCTGTGCGCCAGGCTGGAAGGGCAGCC 1141
Db 4038 ACAGAGACCCCGCGCATCCGCTACCAAGTACAAGCTGTGCGCCAGGCTGGAAGGGCAGCC 4097
QY 1142 CCAGCATCTCCAGCAGCATGACCAAGATCCTGGAGCCCTTCGCGCCCGCAGCCCGC 1201
Db 4098 CCGCCATCTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCGCAAGAGAACCCCG 4157
QY 1202 AGATCGTATCTACAGTACATGGAGGCTGTGAGTGGGCGAGCAGCTGGAGATCGGCC 1261
Db 4158 ACATCGTATCTACAGTACATGGAGGCTGTGAGTGGGCGAGCAGCTGGAGATCGGCC 4217
QY 1262 AGCAGCGGCCAAGATCGAGGAGCTGCGCAAGCACTGCTGGCGTGGGCTTCACCAACC 1321
Db 4218 AGCAGCGGCCAAGATCGAGGAGCTGCGCAAGCACTGCTGGCGTGGGCTTCACCAACC 4277

QY 1322 CCACAAAGAACACCAAGAGAGCCCGCCCTTCTGTGGATGGCTACGAGCTGCACCCCG 1381
Db 4278 CCGACAAGAACACCAAGAGAGCCCGCCCTTCTGTGGATGGCTACGAGCTGCACCCCG 4337
QY 1382 ACAAGTGGACCGTGCAGCCCATCGAGCTGCGCGAGAGAGAGCTGGACCGTGAAGACA 1441
Db 4338 ACAAGTGGACCGTGCAGCCCATCGAGCTGCGCGAGAGAGAGCTGGACCGTGAAGACA 4397
QY 1442 TCCAGAGCTGGTGGCAAGCTGAACCTGGCCAGCAGCATCTACCCGGCATCANGTGC 1501
Db 4398 TCCAGAGCTGGTGGCAAGCTGAACCTGGCCAGCAGCATCTACCCGGCATCANGTGC 4457
QY 1502 GCAGCTGTCAAAGCTGCTGCGCGCGCCCAAGGCCCTGACCGACATCTGCTCCCTGACCC 1561
Db 4458 GCCAGCTGTCAAAGCTGCTGCGCGCGCCCAAGGCCCTGACCGAGGTGGTCCCTGACCC 4517
QY 1562 AGAGCGCGAGCTGGAGCTGGCGCGAGAACCGCGAGATCTCTGCGCGAGCCGCTGACGGCG 1621
Db 4518 AGAGCGCGAGCTGGAGCTGGCGCGAGAACCGCGAGATCTCTGAAAGGAGCCGCTGACGGCG 4577
QY 1622 TGTACTACGACCCACGACAGGACCTGCTGGCCGAGATCCAGAGAGCCGACAGT 1681
Db 4578 TGTACTACGACCCACGACAGGACCTGATCGCCGAGATCCAGAGAGCCGACAGGCGCAGT 4637
QY 1682 GGACCTTACCAGATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCCGCAAGTACGCCAAGA 1741
Db 4638 GGACCTTACCAGATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCCGCAAGTACGCCCGCA 4697
QY 1742 TGGCGACCGCCACACACAGAGCTGAAGAGCTGACCGAGGCCGTGCGAGAAGATCGCCA 1801
Db 4698 TGAAGGCGCCCGCCACACAGAGCTGAAGCAGCTGACCGAGGCCGTGCGAGAAGATCGCCA 4757
QY 1802 TGGAGAGCATCTGATCTGGGCAAGACCCCAAGTTCGCCCTGCCATCCCAAGAGAGA 1861
Db 4758 CCGAGAGCATCTGATCTGGGCAAGACCCCAAGTTCAGCTGCCCCATCCAGAGAGAGA 4817
QY 1862 CTTGGAGACCTTGGTGACCGACTACTGCGAGGCCACCTTGATTCGCCCTGAGTGGAGTTCG 1921
Db 4818 CTTGGAGACCTTGGTGACCGACTACTGCGAGGCCACCTTGATTCGCCCTGAGTGGAGTTCG 4877
QY 1922 TGAACACCCCGCCCTTGGTGAAGCTGTGTTACCAAGCTGGAGAGGAGCCCATCATCGCG 1981
Db 4878 TGAACACCCCGCCCTTGGTGAAGCTGTGTTACCAAGCTGGAGAGGAGCCCATCATCGCG 4937
QY 1982 CCGAGACCTTCTACGTGAGCGCGCGCAACCGCAGACCAAGATCGGCAAGCGCGCT 2041
Db 4938 CCGAGACCTTCTACGTGAGCGCGCGCAACCGCAGACCAAGCTGGGCAAGCGCGCT 4997
QY 2042 AGTGACCGACCGGGCGCGCAGAGATCTGAGCCTGACCGAGACCAACCAAGAGA 2101
Db 4998 ACGTGACCGACCGCGCGCGCAGAGTGTGTCCTGACCGACACCAACCAAGAGA 5057
QY 2102 CCGAGCTGAGGCCATCCAGCTGGCCCTGCGAGGACAGCGGCGAGGCTGAACATCGTGA 2161
Db 5058 CCGAGCTGAGGCCATCCAGCTGGCCCTGCGAGGACAGCGGCTGGAGGTGAACATCGTGA 5117
QY 2162 CCGACAGCCAGTACGCCCTTGGGCATCATCCAGGCCAGCGCGCAAGAGAGGAGCGAGCGAGC 2221
Db 5118 CCGACAGCCAGTACGCCCTTGGGCATCATCCAGGCCAGCGCGCAAGAGAGGAGCGAGC 5177
QY 2222 TGGTGAACCGAGTATCTGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGGTGC 2281
Db 5178 TGGTGAACCGAGTATCTGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGGTGC 5237
QY 2282 CCGCCCAAGGCGCATCGCGCGCAAGCAGCATCGACAAGCTGGTGAAGAGGGCATCC 2341
Db 5238 CCGCCCAAGGCGCATCGCGCGCAAGCAGCAGCTGGAGCGGCTGGTGAAGCGCGGCGATCC 5297
QY 2342 GCAAGGTGCTTCTCTGGAGCGGCATCGATGGGGCATCTGTGATCTACCACTA 2393
Db 5298 GCAAGGTGCTTCTCTGGAGCGGCATCGACAAGGCCCGCAGGAGGAGCAGAGAA 5349

```
RESULT 11
AX427921          9189 bp      DNA      linear      PAT 20-JUN-2002
LOCUS
DEFINITION      Sequence 159 from Patent WO0232943.
ACCESSION       AX427921
VERSION         AX427921.1 GI:21538008
KEYWORDS
SOURCE          synthetic construct.
ORGANISM        artificial sequences.
REFERENCE
AUTHORS         Huang,Y. and Nabel,G.J.
TITLE           Modifications of hiv env, gag, and pol enhance immunogenicity for
                genetic immunization
JOURNAL         Patent: WO 0232943-A 159 25-APR-2002;
                GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source          1. 9189
                /organism="synthetic construct"
                /db_xref="taxon:32630"
                /note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT      2232 a 2671 c 2524 g 1762 t
ORIGIN
Query Match      81.8%; Score 2018.6; DB 6; Length 9189;
Best Local Similarity 91.4%; Pred. No. 1.7e-203;
Matches 2190; Conservative 0; Mismatches 184; Indels 21; Gaps 4;
Qy 14 TGGCCGAGGCATGAGCCAGG---CCACGAGCCCAACATCTTGATGAGCGCAGCAACT 70
Db 2967 TGGCCGAGGCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGAGCGCGCAACT 3026
Qy 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTTCACCTGGCGGCAAGGAGGCCACATCGGCC 130
Db 3027 TCCGCAACCAAGCAAGATCGTGAAGTGTTCACCTGGCGGCAAGGAGGCCACACCGCC 3086
Qy 131 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAGTGGCGGAAGGAGGCCACACAGA 190
Db 3087 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAGTGGCGGAAGGAGGCCACACAGA 3146
Qy 191 TGAAGGACTGCACGAGCGCCAGCGCCAACTTCTCCGCGAGGACCTGGCCCTCCCGCAGG 250
Db 3147 TGAAGGACTGCACGAGCGAGCAGACTAA-----TAGGGAAGATCTGGCCCTTCCCAAG 3200
Qy 251 GCAAGGCGCGGAGTTCGCCAGCGAGCAGACACCGCGCCACAGCGCCACAGCGCGGAGC 310
Db 3201 GGAAGGCGGAGGAATTTCTTCAGAGCAGACGAGCCACAGCGCCACAGCAGAGAGAGC 3260
Qy 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGAGCGCCAGGGCA 364
Db 3261 TTCAGGTTTGGGAAGAGACACAACACTCCCTCTCAGRAGCAGAGCGCATAGACAAGAA 3320
Qy 365 -----CCCTGAATCTCCCGCAGATCACCCCTGTGGCAGCGCCCGCTGTGAGCATCAAG 418
Db 3321 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTGTCTACAATAAGA 3380
Qy 419 TGGCGGCCCATCAAGGAGGCCCTGTGGACACCGCGCGGCGAGCACACCTGTGTGAGG 478
Db 3381 TAGGGGCCAGCTGAAGGAGGCCCTTTAGACACCGCGCGGCGAGCACACCTGTGTGAGG 3440
Qy 479 AGATGAGCTTCCCGGCAAGTGGGAAGCCCAAGATGATCGGGGGCATCGGGCGCTTCATCA 538
Db 3441 AGATGAACCTGCCCGCGCTTGAAGGCCCAAGATGATCGGGCGCATCGGGCGCTTCATCA 3500
Qy 539 AGGTGCGCCAGTACGACCATCTCTGATCGAGATCTCGCGCAAGAAGGCCATCGGCACCG 598
Db 3501 AGGTGGGCCAGTACGACCATCTCTGATCGAGATCTCGGGCCACAAGGCCATCGGCACCG 3560
Qy 599 TGCTGATGCGGCCCGACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCATCGGGCT 658
Db 3561 TGCTGGTGGGCCCGACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCATCGGGCT 3620
Qy 659 GCACCTGAACCTTCCCATCAGCCCATCAGAGACCGTCCCGTGAAGCTGAAGCCCGCA 718
Db 3621 GCACCTGAACCTTCCCATCAGCCCATCAGAGACCGTCCCGTGAAGCTGAAGCCCGCA 3680
Qy 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGAACGAGGAGAGATCAAGCCCTGACCG 778
Db 3681 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGAACGAGGAGAGATCAAGCCCTGAGTG 3740
Qy 779 CCATCTCGAGGAGATGGAAGAGGAGGCAAGATCACCAGATCGGCCCGCCGAGAACCCCT 838
Db 3741 AGATCTGCACCGAGATGGAAGAGGAGGCAAGATCAGCAAGATCGGCCCGCCGAGAACCCCT 3800
Qy 839 ACAACACCCCGCTGTTCGCCATCAAGAAGAAGGACAGCACCACCAAGTGGCGCAAGCTGGTG 898
Db 3801 ACAACACCCCGCTGTTCGCCATCAAGAAGAAGGACAGCACCACCAAGTGGCGCAAGCTGGTG 3860
Qy 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGGAGGTGCAGTGGGCAATCCCC 958
Db 3861 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGGAGGTGCAGTGGGCAATCCCC 3920
Qy 959 ACCCGCGCGCTGAAGAAGAAGAGAGCGTGCACCGTCTGGACGTGGGCGACGCGCTACT 1018
Db 3921 ACCCGCGCGCTGAAGCAGGAAGAGCGTGCACCGTCTGGACGTGGGCGACGCGCTACT 3980
Qy 1019 TCAGCTGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 1078
Db 3981 TCAGCTGCCCTGGACAGGAGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 4040
Qy 1079 ACAACGAGACCCCGCGCATCCGCTACCAAGTACAAGTGTCTGCCCGAGGGCTGGGAAGGCA 1138
Db 4041 ACAACGAGACCCCGCGCATCCGCTACCAAGTACAAGTGTCTGCCCGAGGGCTGGGAAGGCA 4100
Qy 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCTTGAGAGCCCTTCGCGGCCCGCAACC 1198
Db 4101 GCGCGCCAGCATCTTCCAGTGCAGCATGACCAAGATCTTGAGAGCCCTTCGCGCAAGCAACC 4160
Qy 1199 CCGAGATCGTGATCTACCAGTACATGGAGACCTGTAGTGGGAGCGACCTTGGAGATCG 1258
Db 4161 CCGAGATCGTGATCTACCAGTACATGGAGACCTGTAGTGGGAGCGACCTTGGAGATCG 4220
Qy 1259 GCCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGGCTTCAACA 1318
Db 4221 GCCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGGCTTCAACA 4280
Qy 1319 CCGCGCACAAAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1378
Db 4281 CCGCGCACAAAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4340
Qy 1379 CCGCAAGTGGACCGTGCAGCGCATCGAGCTGCCGAGAGGAGGAGGAGGAGGAGGAGGAG 1438
Db 4341 CCGCAAGTGGACCGTGCAGCGCATCGAGCTGCCGAGAGGAGGAGGAGGAGGAGGAGGAG 4400
Qy 1439 ACATCCAGAAGCTGGTGGCAAGCTGAACCTGGCGCAGCAGATCTTACCCCGCATCAAG 1498
Db 4401 ACATCCAGAAGCTGGTGGCAAGCTGAACCTGGCGCAGCAGATCTTACCCCGCATCAAG 4460
Qy 1499 TGGCGCAGCTGTGAAGCTGTGCGCGCGGCGCAAGGCGCTTACCGACATCTGTGCCCTGA 1558
Db 4461 TGGCGCAGCTGTGAAGCTGTGCGCGCGGCGCAAGGCGCTTACCGACATCTGTGCCCTGA 4520
Qy 1559 CCGAGGAGGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGCGCGAGGCGCGTGCAGG 1618
Db 4521 CCGAGGAGGCGGAGCTGGAGCTGGCGGAGAACCGCGGAGATCTGCGCGAGGCGCGTGCAGG 4580
Qy 1619 GCGTGTACTACGACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1678
Db 4581 GCGTGTACTACGACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4640
Qy 1679 AGTGGACCTTACAGATCTTACAGGAGGCGCTTCAAGAACTTCAAGACCGCGCAAGTACGCCA 1738
Db 4641 AGTGGACCTTACAGATCTTACAGGAGGCGCTTCAAGAACTTCAAGACCGCGCAAGTACGCC 4700
Qy 1739 AGATGCGCACCGCCACACCAACGACGTGAAGCAGCTGACCGAGCGCGTGCAGAGATCG 1798
```


QY 479 AGATGAGCTGCGCGCAAGTGAAGCCCAAGATGATCGCGGCATCGCGGCTTCATCA 538
DB 3441 AGATGAACCTGCGCGCGCGCTGGAAGCCCAAGATGATCGCGGCATCGCGGCTTCATCA 3500
QY 539 AGGTGGCCAGTACGACCAAGATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
DB 3501 AGGTGGCCAGTACGACCAAGATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 3560
QY 599 TGTGATCGGCGCCACCGCCGCTGAACATCATCGGCGGCAACATGCTGACCCAGCTGGGCT 658
DB 3561 TGTGATCGGCGCCACCGCCGCTGAACATCATCGGCGGCAACATGCTGACCCAGCTGGGCT 3620
QY 659 GCACCTGAACCTTCCCATACGCGCATCGAGCCGCTGCGCGGCAAGAGCTGAAGCCGCGCA 718
DB 3621 GCACCTGAACCTTCCCATACGCGCATCGAGCCGCTGCGCGGCAAGAGCTGAAGCCGCGCA 3680
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGAAGCCCTGACCGAGGAGAGATCAAGGCGCTGACCG 778
DB 3681 TGGACGCGCCCAAGGTGAAGCAGTGAAGCCCTGACCGAGGAGAGATCAAGGCGCTGAGTG 3740
QY 779 CCATCTGCGAGAGATGGAAGAGGAGGCAAGATCAACCAAGATCGGCGCCGCGAGAACCCCT 838
DB 3741 AGATCTGCAACCGAGATGGAAGAGGAGGCAAGATCAACCAAGATCGGCGCCGCGAGAACCCCT 3800
QY 839 ACAACACCCCGCTGTCGCATCAAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 898
DB 3801 ACAACACCCCGCTGTCGCATCAAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3860
QY 899 ACTTCCGCGAGCTGAACAGCGCACCGAGGCTTCTGGGAGGTGCAAGTGGGCGCTGCGCC 958
DB 3861 ACTTCCGCGAGCTGAACAGCGCACCGAGGCTTCTGGGAGGTGCAAGTGGGCGCTGCGCC 3920
QY 959 ACCCGCGCGCTGGAAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1018
DB 3921 ACCCGCGCGCTGGAAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3980
QY 1019 TCAGCTGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1078
DB 3981 TCAGCTGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4040
QY 1079 ACAACGAGACCGCGCGCATCCGCTACCAAGTACCAAGTGTGCGCGGAGGAGGAGGAGG 1138
DB 4041 ACAACGAGACCGCGCGCATCCGCTACCAAGTGTGCGCGGAGGAGGAGGAGGAGGAGG 4100
QY 1139 GCGGAGCATCTCCAGAGCAGCATGACCAAGATCTGAGGCGCTTCCCGCGCGCAAGC 1198
DB 4101 GCGGAGCATCTCCAGAGCAGCATGACCAAGATCTGAGGCGCTTCCCGCGCGCAAGC 4160
QY 1199 CGGAGTCTGTATCTACCAAGTACATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1258
DB 4161 CGGAGTCTGTATCTACCAAGTACATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4220
QY 1259 GCGAGACCGCGCGCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1318
DB 4221 GCGAGACCGCGCGCATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4280
QY 1319 CCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1378
DB 4281 CCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4340
QY 1379 CGGACAAGTGGAGCGGCGCATCGAGCTGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGG 1438
DB 4341 CGGACAAGTGGAGCGGCGCATCGAGCTGCGCGGAGAGGAGGAGGAGGAGGAGGAGGAG 4400
QY 1439 ACATCAGAGCTGGTGGGAGAGCTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1498
DB 4401 ACATCAGAGCTGGTGGGAGAGCTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4460
QY 1499 TGGCGGAGCTGCAAGCTGCTGCGCGGCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1558
DB 4461 TGGCGGAGCTGCAAGCTGCTGCGCGGCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4520

QY 1559 CCGAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTCGCGAGCGCGCTGCACG 1618
DB 4521 CCGAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGAAGGAGCGCGCTGCACG 4580
QY 1619 GCGTGTACTACGACCCCGACCAAGAGGAGCTGGTGGCGGAGATCCAGAGAGGAGGCGCACGAC 1678
DB 4581 GCGTGTACTACGACCCCGACCAAGAGGAGCTGATCGCGGAGATCCAGAGAGGAGGCGCACGAC 4640
QY 1679 AGTGGACCTTACAGATCTTACAGGAGCGCTTCAAGAACCTTGAGAGAGGAGGCGCACGAC 1738
DB 4641 AGTGGACCTTACAGATCTTACAGGAGCGCTTCAAGAACCTTGAGAGAGGAGGCGCACGAC 4700
QY 1739 AGATGGCGACCGCGCGACCAACAGAGCTGAAGCAGCTGACCGAGGCGCTGACAGAGATCG 1798
DB 4701 GCATGAGGCGCGCGCGACCAACAGAGCTGAAGCAGCTGACCGAGGCGCTGACAGAGATCG 4760
QY 1799 CCATGAGAGGAGCTGATGATCTGGGGCAAGACCCCAAGTTCGCCCTTGCCTCATCAGAGG 1858
DB 4761 CCATGAGAGGAGCTGATGATCTGGGGCAAGACCCCAAGTTCGCCCTTGCCTCATCAGAGG 4820
QY 1859 AGACCTGGGAGAGCTGGTGGACCGGACTACTGGCAGGCGCACTGGATCCCGAGTGGGAGT 1918
DB 4821 AGACCTGGGAGAGCTGGTGGACCGGACTACTGGCAGGCGCACTGGATCCCGAGTGGGAGT 4880
QY 1919 TCGTGAACACCGCGCGCGCTGGTGAAGCTGTGTACAGCTGGAGAGGAGGAGGAGGAGGAGG 1978
DB 4881 TCGTGAACACCGCGCGCGCTGGTGAAGCTGTGTACAGCTGGAGAGGAGGAGGAGGAGGAG 4940
QY 1979 GCGCGGAGACCTTCTACGTGGAGCGCGCGCAACCGCGAGAGGAGGAGGAGGAGGAGGAGG 2038
DB 4941 GCGCGGAGACCTTCTACGTGGAGCGCGCGCAACCGCGAGAGGAGGAGGAGGAGGAGGAGG 5000
QY 2039 GCTACGTGACCGCGCGCGCGCGCAAGATCTGAGAGCTGACCGAGAGGAGGAGGAGGAGGAG 2098
DB 5001 GCTACGTGACCGCGCGCGCGCGCAAGATCTGAGAGCTGACCGAGAGGAGGAGGAGGAGGAG 5060
QY 2099 AGACCGAGCTGAGGCGCATCCAGCTGGCGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 2158
DB 5061 AGACCGAGCTGAGGCGCATCCAGCTGGCGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAG 5120
QY 2159 TGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2218
DB 5121 TGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5180
QY 2219 AGCTGTGTGAACAGAGATCATCGAGCAGCTGATCAAGAGAGGAGGAGGAGGAGGAGGAGG 2278
DB 5181 AGCTGTGTGAACAGAGATCATCGAGCAGCTGATCAAGAGAGGAGGAGGAGGAGGAGGAGG 5240
QY 2279 TGCG 2338
DB 5241 TGCG 5300
QY 2339 TCGCGAGGAGTGTGTTCTTGGAGGCGCATCGATGGCGGCGCATCGTGTGATCTACCACTA 2393
DB 5301 TCGCGAGGAGTGTGTTCTTGGAGGCGCATCGATGGCGGCGCATCGTGTGATCTACCACTA 5355

RESULT 14
AX427935

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AX427935
Sequence 173 from Patent W00232943.
AX427935
AX427935.1 GI:21538022

synthetic construct.
synthetic construct
artificial sequences.

1

Huang, Y. and Nabel, G. J.
Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
Patent: WO 0232943-A 173 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)

9782 bp

DNA

linear

PAT 20-JUN-2002

FEATURES	source	Location/Qualifiers
	1. .9782	
	/organism="synthetic construct"	
	/db_xref="taxon:32630"	
	/note="plasmid pVR1012x/s containing HIV genes"	
BASE COUNT	2339 a 2895 c 2719 g 1829 t	
ORIGIN		
Query Match	81.7%;	Score 2017; DB 6; Length 9782;
Best Local Similarity	91.4%;	Pred. No. 2.5e-203;
Matches 2189;	Conservative 0;	Mismatches 185; Indels 21; Gaps 4;
Qy 14	TGGCGGAGCCATGAGCCAGG---CCACAGCGCCAAACATCCTGTATGAGCGGACGCAACT 70	
Db	TGGCGGAGCCATGAGCCAGGCTGACACAGCGCCACCACCATGATGATGAGCGGCAACT 3638	
Qy 71	TCAAGGGCCCCAAGCGCATCATCAAGTCTTCACTTCCGCGGCAAGAGGGCCACATCGGCC 130	
Db	TCCGCAACACCGCAAGATCGTGAAGTCTTCAACTTGGCGGAAGGAGGGCCACACCGCCC 3698	
Qy 131	GCAACTCGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACCAGA 190	
Db	GCAACTCGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACCAGA 3758	
Qy 191	TGAAGGACTGCACGAGCGCCAGGCCAACTTCTTCCCGGAGGACCTTGCCTTCCCCCAGG 250	
Db	TGAAGGACTGCACGAGCGGACAGCGTAA-----TAGGGAAGATCTGGCCTTCCCAACAAG 3812	
Qy 251	GCAAGGCCCGGAGTTCCCGAGCGAGCAGAACCGCGCCCAACAGCCCCACCGCCGAGCG 310	
Db	GGAAGGCCAGGGAAATTTTCTTCAGAGCAGACACAGCCAGCCAAACAGCCCCACCAAGAGAGCG 3872	
Qy 311	TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCCGAGCGCCAGGCA 364	
Db	TTCAGGTTTGGGAAGAGACAACTCCCTCTCAGAGCAGGAGCCCATAGACRAGGAA 3932	
Qy 365	-----CCCTGAATTTCCCCCAGATFACCTCTGTGGCAGCGCCCTGTGTGAGCATCAAGG 418	
Db	CTGTATCCTTTAGCTTCCTCAGATCACTTTTGGCAGCGACCCCTCGTCAACAATAAGA 3992	
Qy 419	TGGCGGCGCAGATCAAGGAGCGCCTGCTGGACACGGCGCGCAGACACACCTGTGTGGAGG 478	
Db	TAGGGGGCCACTGAAGGAGCGCCTTCTAGACACGGCGCGCGCAGACACCTGTGTGGAGG 4052	
Qy 479	AGATGAGCCTGCCCGGAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538	
Db	AGATGAACCTGCCCGCGCGCTGGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 4112	
Qy 539	AGGTGCGCCATGACACCATGATCTGTATGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCG 598	
Db	AGGTGGGCGCATGAGCACCAGATCTGTATCGAGATCTGCGGCAACAAGGCCATCGGCACCG 4172	
Qy 599	TGCTGATCGGGCCACACCCCGTGAACATCATCGCCCGCAAGCTGCCCTGAAGCTGAAGCCCGGCA 718	
Db	TGCTGTGGGGCCACCCCGGTGAACATCATCGGCGCGCAAGCTGCTGACCCAGATCGGCT 4232	
Qy 659	GCACCTCTGAATTTCCCATAGCCCCCATCGAGACCGTGCCTGTGAAGCTGAAGCCCGGCA 718	
Db	GCACCTCTGAATTTCCCATAGCCCCCATCGAGACCGTGCCTGTGAAGCTGAAGCCCGGCA 4292	
Qy 719	TGGAGGCCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAAGATCAAGGCCCTTGACCG 778	
Db	TGGAGGCCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAAGATCAAGGCCCTTGTGG 4352	
Qy 779	CCATCTCGAGGAGATGAGAGAGGAGGCAAGATCCCAAGATCGGCCCCCGAGACCCCT 838	
Db	AGATCTGCAGATGAGAGAGGAGGCAAGATCAGCAAGATCGGCCCCCGAGAACCCCT 4412	
Qy 839	ACAACACCCCGTGTTCGCCATCAAGAGAAGAGCAGCACCAAGTGGCGCAAGCTGGTGG 898	
Db	ACAACACCCCGTGTTCGCCATCAAGAGAAGAGCAGCACCAAGTGGCGCAAGCTGGTGG 4472	
Qy 899	ACTTCCGGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTCAGCTGGGCATCCCCC 958	
Qy 4473	ACTTCCGGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTCAGCTGGGCATCCCCC 4533	
Qy 959	ACCCCGCGGCTTGAAGAGAAGAGAGCGTGACCGTGTGACGTGGGAGCGGCGCTACT 1018	
Db	ACCCCGCGGCTTGAAGAGAAGAGAGCGTGACCGTGTGACGTGGGAGCGGCGCTACT 4592	
Qy 1019	TCAGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCTTACACCTCCCGAGCATCA 1078	
Db	TCAGCGTGCCTTGGACAGGACTTCCGCAAGTACACCGCTTACACCTCCCGAGCATCA 4552	
Qy 1079	ACAAGAGACCCCGGCTTCCAGAGTACAGTACAACTGTGTCGCCCGAGGCTGGAAGGGCA 1138	
Db	ACAAGAGACCCCGGCTTCCAGAGTACAGTACAACTGTGTCGCCCGAGGCTGGAAGGGCA 4712	
Qy 1139	GCCCGAGCATTTTCAGAGCAGCATGACCAAGATCCTGAGAGCCCTTCCGGGCCCGCAACC 1198	
Db	GCCCGAGCATTTTCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGAACCC 4772	
Qy 1199	CCGAGATCGTGATCTACCACTATCATGAGAGCTGTACGTGGGACGAGCTCTGGAGATCG 1258	
Db	CCGACATCGTGATCTACCACTATCATGAGACACCTGTACGTGGGACGAGCTCTGGAGATCG 4832	
Qy 1259	GCCAGACCGCGCCCAAGATCGAGGAGCTGCGCAAGCACCTGTGTCGCTGGGGCTTCCACCA 1318	
Db	GCCAGACCGCGCACCAAGATCGAGGAGCTGCGCGCAGCACCTGCTGCGCTGGGGCTTCCACCA 4892	
Qy 1319	CCCCCGACAGAGAACACAGAGGAGCGCCCTTCTGTGTGGATGGGCTTACGAGCTGCACC 1378	
Db	CCCCCGACAAGAGAACCAAGAGGAGCGCCCTTCTGTGTGGATGGGCTTACGAGCTGCACC 4952	
Qy 1379	CCGACAAGTGGACCGTGCAGCGCATCGAGCTGCCGAGAGGAGAGCTGCACCGTGAACG 1438	
Db	CCGACAAGTGGACCGTGCAGCGCATCGAGCCCTCTGCCGAGAGGAGACCTGCACCGTGAACG 5012	
Qy 1439	ACATCCAGAAAGCTGTGTGGSCAAAGCTGAACCTGGGCGCCAGCCAGATCTACCCCGGCATCAAGG 1498	
Db	ACATCCAGAAAGCTGTGTGGSCAAAGCTGAACCTGGGCGCCAGCCAGATCTACCGCGGCATCAAGG 5072	
Qy 1499	TGCGCCAGCTGTGCAAGCTGTGCGGGGGGCCAAGGCCCTGACCGACATCTGTGCCCTTGA 1558	
Db	TGCGCCAGCTGTGCAAGCTGTGCGGGGGGCCAAGGCCCTGACCGAGCGCCGTCGACG 5132	
Qy 1559	CCGAGAGGCCAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGCGCCGTCGACG 1618	
Db	CCGAGAGGCCAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGAAGAGACCCCGTCGACG 5192	
Qy 1619	GGGTGCTACTACGACCCCAAGAGACCTTGGTGGCGGAGATCCAGAGAGCGGCGCACGACC 1678	
Db	GGGTGCTACTACGACCCCAAGAGACCTTATCGCGAGATCCAGAGAGCGGCGCACGACC 5252	
Qy 1679	AGTGACCTACCATGATCTACAGGAGGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1738	
Db	AGTGACCTACCATGATCTACAGGAGGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCC 5312	
Qy 1739	AGATCGGACCCCGCCACACCAACAGCTGACAGCTGACGAGGCGCGTGCAGAAAGTACG 1798	
Db	AGATCGGACCCCGCCACACCAACAGCTGACAGGAGCGCGTGCAGAAAGTACG 5372	
Qy 1799	CAATGGAGAGCATCTGTATCTGGGGAAGACCCCAAGTTCGCCCTGCCCATCCAGAGG 1858	
Db	CAATGGAGAGCATCTGTATCTGGGGAAGACCCCAAGTTCAGAGCTGCCCATCCAGAGG 5432	
Qy 1859	AGACCTGGAGACCTGTGTGGACCGCTACTTGCAGCGCACCTTGGATCCCCCGAGTGGGAGT 1918	
Db	AGACCTGGAGGCGCTGGTGGACCGGAGTACTGGCAGGCCACCTTGGATCCCCCGAGTGGGAGT 5492	
Qy 1919	TCGTGAACACCCCGCTTGGTGAAGCTGTGTATACAGCTGGAGAGGAGGCCCATCATCG 1978	
Db	TCGTGAACACCCCGCTTGGTGAAGCTGTGTATACAGCTGGAGAGGAGGCCCATCATCG 5552	
Qy 1979	GGCGCGAGACCTTCTACGTGGACGGCGCGCAACCGCGAGACCAAGATTCGCAAGGCCG 2038	

Db	5553	CGCGCGAGACCTTCTACGTGGACGGCGCGGCCCAACCGCGGAGACCAAGCTGGGCAAGGCCG	5612
Qy	2039	GCTACGTGACCGACCGGGGCGCGCAGAAGATCGTGTAGCGCTCACCGGAGACCAACCAACAG	2098
Db	5613	GCTACGTGACCGACCGGGGCGCGCAGAAGTGTGTGCCCTGACCGACACCAACCAACAG	5672
Qy	2099	AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGSCAGGACAGCGCGGACGAGTGAACATCG	2158
Db	5673	AGACCGAGCTGCAGGCCATCCACCTGGCCCTGCGAGGACAGCGGCTGGAGGTGAACATCG	5732
Qy	2159	TGACCGACAGCCAGTAGCGCCCTGGGCATCATCCAGGCGCCAGCCGACAGAGCGAGAGCG	2218
Db	5733	TGACCGACAGCCAGTAGCGCCCTGGGCATCATCCAGGCGCCAGCCGACAGAGCGAGAGCG	5792
Qy	2219	AGCTGGTGAACACAGATCATCGACGAGCTGTATCAAGAGGAGAAGGTGTACCTGAGCTGGG	2278
Db	5793	AGCTGGTGAACACAGATCATCGACGAGCTGTATCAAGAGGAGAAGGTGTACCTGAGCTGGG	5852
Qy	2279	TGCCCGCCCAAGAGGCATCGCGGCAACGACGAGATCGACAAGCTGGTGAGCAAGGGCA	2338
Db	5853	TGCCCGCCCAAGAGGCATCGCGGCAACGACGAGGTGGAGCGGCTGGTGAGCGCGGCA	5912
Qy	2339	TCCGCAAGGTGCTTCTCTGACGCGCATCGATGCGGCATCGTATCTACCACTA	2393
Db	5913	TCCGCAAGGTGCTTCTCTGACGCGCATCGACAAGGCGCCAGGAGGAGCACAGAA	5967
RESULT 15			
AX427934	AX427934	Sequence 172 from Patent WO0232943.	linear PAT 20-JUN-2002
LOCUS	AX427934	9783 bp DNA	
DEFINITION	Sequence 172 from Patent WO0232943.		
ACCESSION	AX427934		
VERSION	AX427934.1	GI:21538021	
KEYWORDS	synthetic construct.		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1		
AUTHORS	Huang Y. and Nabel G.J.		
TITLE	Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization		
JOURNAL	Patent: WO 0232943-A 172 25-APR-2002;		
FEATURES	GOVERNMENT OF THE UNITED STATES (US)		
source	Location/Qualifiers		
	1..9783		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	/note="plasmid pVR1012x/s containing HIV genes"		
BASE COUNT	2339 a 2897 c 2723 g 1824 t		
ORIGIN			
Query Match 81.7%; Score 2017; DB 6; Length 9783;			
Best Local Similarity 91.4%; Pred. No. 2.5e-203;			
Matches 2189; Conservative 0; Mismatches 185; Indels 21; Gaps 4;			
Qy	14	TGGCGAGGCCATGAGCCAGG---CCACGACGGCCAAACATCTGTATGAGCGGACGCAACT	70
Db	2964	TGGCGAGGCCATGAGCCAGGTGACCAACACGCGCCACCATCATGATGAGCGCGGCAACT	3023
Qy	71	TCAAGGGCCCCAAGCGCATCATCAAGTGCTTTCAACTGCGCAAGGAGGGCCACATCGGCC	130
Db	3024	TCCGCAACCGCGCAAGATCGTGAAGTGCTTCAACTGCGCAAGGAGGGCCACACCGGCC	3083
Qy	131	GCACTGCGCGCCCCCGCGAAGAGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA	190
Db	3084	GCAACTGCGCGCCCCCGCGAAGAGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA	3143
Qy	191	TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTCCGGGAGGAGCTGGCCCTCCCCAGG	250
Db	3144	TGAAGGACTGCACCGAGCGCCAGGCTAA-----TAGGGAAGATCTGGCCCTCCACAG	3197
Qy	251	GCAAGGCCCGCGAGTTCCCCACGCGAGAGAACCGCGCAACAGGCCCCACCAAGCGCGGAG	310

...

Search completed: February 10, 2003, 17:55:41
Job time : 6389.22 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 10:51:03 ; Search time 492.196 Seconds
(without alignments)
11296.684 Million cell updates/sec

Title: US-09-610-313-30

Perfect score: 2469

Sequence: 1 gtcgacgcaccatggccga.....ggcgtagcaccggtagtc 2469

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_l01002:*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2469	100.0	2469	24	ABL39959
2	2442.2	98.9	2463	24	ABL39960
3	2415.4	97.8	2457	24	ABL39961
4	2052	83.1	2312	21	AAA70481
5	2051.2	83.1	9166	24	ABK91616
6	2042.8	82.7	9788	24	ABK91622
7	2028.8	82.2	9169	24	ABK91617
8	2027	82.1	9194	24	ABK91611
9	2027	82.1	12411	24	ABK91613

10	2025.4	82.0	9194	24	ABK91612
11	2025.2	82.0	2306	21	AAA70479
12	2020.4	81.8	9785	24	ABK91624
13	2018.6	81.8	9189	24	ABK91607
14	2017	81.7	9187	24	ABK91619
15	2017	81.7	9170	24	ABK91614
16	2017	81.7	9782	24	ABK91621
17	2017	81.7	9783	24	ABK91620
18	2017	81.7	9792	24	ABK91618
19	1999	80.9	3015	24	ABL39983
20	1998.4	80.9	2300	21	AAA70480
21	1982.4	80.3	3009	24	ABL40024
22	1969.4	79.8	4319	21	AAA70414
23	1935.8	78.4	8908	21	AA252055
24	1907	77.2	2577	24	AAU41592
25	1905.8	77.2	4343	24	AA167896
26	1901.8	77.0	4341	24	AA167897
27	1881.2	76.2	1978	24	ABL39991
28	1879.2	76.1	7897	24	ABK91609
29	1858	75.3	2305	21	AAA70477
30	1837	74.4	1965	24	ABL39989
31	1831.2	74.2	2299	21	AAA70478
32	1815.6	73.5	3012	21	AA252054
33	1807.6	73.2	9407	24	ABK91623
34	1659.4	67.2	4307	20	AA208740
35	1659.4	67.2	4307	21	AAA93972
36	1659.4	67.2	4307	22	AAC86876
37	1651.4	66.9	4307	22	AAH43696
38	1651.4	66.9	4327	21	AAA93984
39	1651.4	66.9	4353	21	AAA93983
40	1651.4	66.9	4642	21	AAA93982
41	1651.4	66.9	9772	22	AAH43699
42	1608	65.1	1680	24	ABL39997
43	1592.8	64.5	4338	22	AAH22806
44	1592.8	64.5	8366	22	AAH22810
45	1554.4	63.0	1668	24	ABL39999

ALIGNMENTS

RESULT 1
ABL39959
ID ABL39959 standard; DNA; 2469 BP.
XX
AC ABL39959;
XX
XX 15-MAY-2002 (first entry)
XX Synthetic construct PR975(+) SEQ ID NO:30.
XX Human immunodeficiency virus type C; antigenic HIV type C protein;
XX immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
XX immunostimulant; gene therapy; gene; ds.
XX Human immunodeficiency virus type C.
OS Synthetic.
XX WO200204493-A2.
XX 17-JAN-2002.
XX
XX 05-JUL-2001; 2001WO-US21241.
XX
XX 05-JUL-2000; 2000US-0610313.
XX (CHIR) CHIRON CORP.
XX (UYST-) UNIV STELLENBOSCH.
XX
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX WPI; 2002-154920/20.
XX

Modified HIV prote
HIV FS(-)_ProtMod_
Modified HIV prote
Modified HIV prote
Modified HIV prote
Modified HIV prote
Modified HIV prote
Modified HIV prote
Synthetic Pol poly
HIV FS(-)_ProtMod_
Synthetic Pol poly
HIV Gag-polymerase
Packaging construc
HIV-1 subtype C is
Nucleotide sequenc
Nucleotide sequenc
Synthetic protease
Modified HIV prote
HIV FS(+)_Protinac
Synthetic protease
HIV FS(+)_Protinac
Codon optimised Hu
Modified HIV prote
HIV gagpol-syNp c
Human immunodefici
Nucleotide sequenc
Codon optimised HI
HIV partial leader
HIV partial leader
HIV complete leade
Codon optimised ga
Synthetic RT polyn
DNA sequence of a
Synthetic RT polyn

PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
PT cell lines, particularly in gene therapy -

XX Claim 1; Fig 8; 233pp; English.

CC The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
CC Nef (1). (1) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV Type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the
CC exemplification of the present invention.

XX Sequence 2469 BP; 571 A; 833 C; 761 G; 304 T; 0 other;

SQ Query Match 100.0%; Score 2469; DB 24; Length 2469;
Best Local Similarity 100.0%; Pred No. 8, 2e-301;
Matches 2469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGAGCGCCACCATGCGCGAGGCCATGAGCCAGGCCACCAAGCTTCACTGCGCGCAGGAGGGC 120
DB 1 GTCGAGCCACCATGCGCGAGGCCATGAGCCAGGCCACCAAGCTTCACTGCGCGCAGGAGGGC 60
QY 61 CGCAGCAACTTCAAGGCCCCCAAGCGCATCATCAAGTGTTCACCTGCGCGCAGGAGGGC 120
DB 61 CGCAGCAACTTCAAGGCCCCCAAGCGCATCATCAAGTGTTCACCTGCGCGCAGGAGGGC 120
QY 121 CATCATGCGCGCAACTTCCCGCGCCCCCGCAAGAGGGCTGCTGGAAGTCCGCGCAAGGAG 180
DB 121 CATCATGCGCGCAACTTCCCGCGCCCCCGCAAGAGGGCTGCTGGAAGTCCGCGCAAGGAG 180
QY 181 GCGCACAGATGAAGACTGCAACGAGCGCCAGGCCAATCTTCCGCGAGGACCTTGCC 240
DB 181 GCGCACAGATGAAGACTGCAACGAGCGCCAGGCCAATCTTCCGCGAGGACCTTGCC 240
QY 241 TTCCCGCAGGGCAAGCGCGCGAGTTCCCGCAGCGAGCAAGCCGCGCCACAGCCCCCACC 300
DB 241 TTCCCGCAGGGCAAGCGCGCGAGTTCCCGCAGCGAGCAAGCCGCGCCACAGCCCCCACC 300
QY 301 AGCCGCGAGCTGCAAGTGGCGCGCACACCCCGCAGCGAGCCGCGCGCGAGCGCCAG 360
DB 301 AGCCGCGAGCTGCAAGTGGCGCGCACACCCCGCAGCGAGCCGCGCGCGAGCGCCAG 360
QY 361 GGCACCTGAACCTTCCCGCAGATCACCTGTGCGACGCGCCCTGTGTGAGCATCAAGGTG 420
DB 361 GGCACCTGAACCTTCCCGCAGATCACCTGTGCGACGCGCCCTGTGTGAGCATCAAGGTG 420
QY 421 GCGCGCCAGATCAAGAGCGCCCTGTGACACCGCGCGCGAGCACCCTGTGTGAGGAG 480
DB 421 GCGCGCCAGATCAAGAGCGCCCTGTGACACCGCGCGCGAGCACCCTGTGTGAGGAG 480
QY 481 ATGAGCCTGCGCGCAAGTGGAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCAAG 540
DB 481 ATGAGCCTGCGCGCAAGTGGAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCAAG 540
QY 541 GTGCGCCAGTACACAGATCTCTGATGATGATGATGCGGCAAGAGGCCATCGCGCAGGTG 600
DB 541 GTGCGCCAGTACACAGATCTCTGATGATGATGATGCGGCAAGAGGCCATCGCGCAGGTG 600
QY 601 CTGATCGCGCCACCCCTGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCTGC 660
DB 601 CTGATCGCGCCACCCCTGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCTGC 660
QY 661 ACCCTGAACCTTCCCATCATCGCGCCATCGAGACCGTGCCTGTAAGCTGAAGCCCGGCATG 720
DB 661 ACCCTGAACCTTCCCATCATCGCGCCATCGAGACCGTGCCTGTAAGCTGAAGCCCGGCATG 720
QY 721 GACGGCCCCAAGTGAAGCAGTGGCCCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCC 780

DB 721 GACGGCCCCAAGTGAAGCAGTGGCCCCCTGACCGAGGAGATCAAGGCCCTGACCGCC 780
QY 781 ATCTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAGATGCGCCCCCGAGAGACCCCTAC 840
DB 781 ATCTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAGATGCGCCCCCGAGAGACCCCTAC 840
QY 841 AACACCCCGCTGTTCGCCATCAAGAAGAAGAGAGACCAAGTGGCGCAAGCTGGTGGAC 900
DB 841 AACACCCCGCTGTTCGCCATCAAGAAGAAGAGAGACCAAGTGGCGCAAGCTGGTGGAC 900
QY 901 TTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGAGCTGGGCGATCCCCCAG 960
DB 901 TTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGAGCTGGGCGATCCCCCAG 960
QY 961 CCCGCGCGCTGAAGAAGAAGAGAGGCTGACCGTGTGAGCTGGGCGAGCGCTACTTC 1020
DB 961 CCCGCGCGCTGAAGAAGAAGAGAGGCTGACCGTGTGAGCTGGGCGAGCGCTACTTC 1020
QY 1021 AGCGTGGCCCCGTGACGAGGACTTCCGCAAGTACACCGCTTACCATCCCCCAGCATCAAC 1080
DB 1021 AGCGTGGCCCCGTGACGAGGACTTCCGCAAGTACACCGCTTACCATCCCCCAGCATCAAC 1080
QY 1081 AACGAGACCCCGCGCATCCGCTACAGTACACGTGCTGCCCGCAGGGCTGGAAGGGCAGC 1140
DB 1081 AACGAGACCCCGCGCATCCGCTACAGTACACGTGCTGCCCGCAGGGCTGGAAGGGCAGC 1140
QY 1141 CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCGCGCAACCC 1200
DB 1141 CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCGCGCAACCC 1200
QY 1201 GAGATCGTATCTACCAAGTACATGGAGCTGTACCTGGCGAGCGACCTGGAGATCGGC 1260
DB 1201 GAGATCGTATCTACCAAGTACATGGAGCTGTACCTGGCGAGCGACCTGGAGATCGGC 1260
QY 1261 CAGCACCGCGCCAAAGTTCGAGGAGCTGCGCAAGCACCCTGCTGCGCTGGGGCTTCAACACC 1320
DB 1261 CAGCACCGCGCCAAAGTTCGAGGAGCTGCGCAAGCACCCTGCTGCGCTGGGGCTTCAACACC 1320
QY 1321 CCCGCAAGAAGCACCAGAGGAGCCCGCTTCTGATGGGCTACGAGCTGCACCC 1380
DB 1321 CCCGCAAGAAGCACCAGAGGAGCCCGCTTCTGATGGGCTACGAGCTGCACCC 1380
QY 1381 GACAAGTGAACCTGAGCCATCGAGCTGCGCGAGAGAGAGCTGAGCCGTGAACGAC 1440
DB 1381 GACAAGTGAACCTGAGCCATCGAGCTGCGCGAGAGAGAGCTGAGCCGTGAACGAC 1440
QY 1441 ATCCAGAAGCTGTGGGCAAGCTGAACCTGGGCGAGCGAGATCTACCCCGGATCAAGGTG 1500
DB 1441 ATCCAGAAGCTGTGGGCAAGCTGAACCTGGGCGAGCGAGATCTACCCCGGATCAAGGTG 1500
QY 1501 CGCCAGCTGTGAAGCTGTGCGCGCGCGCAAGGCCCTGACCGACATCTGCCCCCTGACC 1560
DB 1501 CGCCAGCTGTGAAGCTGTGCGCGCGCGCAAGGCCCTGACCGACATCTGCCCCCTGACC 1560
QY 1561 GAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGCGCGCTGCACGGC 1620
DB 1561 GAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGCGCGCTGCACGGC 1620
QY 1621 GTGTACTAGACCCCCAGAGACCTGGTGGCGGAGATCCAGAGAGAGGCGCCACGACCGAG 1680
DB 1621 GTGTACTAGACCCCCAGAGACCTGGTGGCGGAGATCCAGAGAGAGGCGCCACGACCGAG 1680
QY 1681 TGGACCTACCATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCTGAGGCAAGTACGCCAAG 1740
DB 1681 TGGACCTACCATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCTGAGGCAAGTACGCCAAG 1740
QY 1741 ATGCGCACCGCCACACCAACGAGCTGAAGAGCTGACCGAGCGCTGCAAGAGATCGCC 1800
DB 1741 ATGCGCACCGCCACACCAACGAGCTGAAGAGCTGACCGAGCGCTGCAAGAGATCGCC 1800
QY 1801 ATGCGAGAGATCTGTATCTGGGCGCAAGACCCCAAGTTCGCGCTGCCCCATTCAGAGAGGAG 1860
DB 1801 ATGCGAGAGATCTGTATCTGGGCGCAAGACCCCAAGTTCGCGCTGCCCCATTCAGAGAGGAG 1860

Db 1801 ATGAGAGCATCGTGTATGCGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAAGGAG 1860
Qy 1861 ACCTGGAGACTGTGACCGACTACTGCGAGGCCACTGTGATCCCGAGTGGAGTTC 1920
Db 1861 ACCTGGAGACTGTGACCGACTACTGCGAGGCCACTGTGATCCCGAGTGGAGTTC 1920
Qy 1921 GTGAACACCCCCCTGTGTAAGCTGTGGTACCAAGCTGGAGAGGCCATCATCGGC 1980
Db 1921 GTGAACACCCCCCTGTGTAAGCTGTGGTACCAAGCTGGAGAGGCCATCATCGGC 1980
Qy 1981 GCCGAGACCTTCTAGCTGGAGGCCGCCCAACCCGAGACCAAGATCGGAAGCCCGGC 2040
Db 1981 GCCGAGACCTTCTAGCTGGAGGCCGCCCAACCCGAGACCAAGATCGGAAGCCCGGC 2040
Qy 2041 TACGTGACCGACCGGGCGCGCAGAAAGATCGTGAGCCCTGACGAGACCAACCAAGAG 2100
Db 2041 TACGTGACCGACCGGGCGCGCAGAAAGATCGTGAGCCCTGACGAGACCAACCAAGAG 2100
Qy 2101 ACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACACGCGCAGCGAGGTGAACATCGTG 2160
Db 2101 ACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACACGCGCAGCGAGGTGAACATCGTG 2160
Qy 2161 ACCGACAGCAGTAGCCCTGGCATCATCCAGGCCATCCAGGCCACGAGCGAGCGAG 2220
Db 2161 ACCGACAGCAGTAGCCCTGGCATCATCCAGGCCATCCAGGCCACGAGCGAGCGAG 2220
Qy 2221 CTGTTGAACACAGATCATCGACAGCTGATCAAGAGGAGAGGTGACTGAGCTGGGTG 2280
Db 2221 CTGTTGAACACAGATCATCGACAGCTGATCAAGAGGAGAGGTGACTGAGCTGGGTG 2280
Qy 2281 CCGGCCCAACAGGGCATCGCGGCAACGAGCAGATCGACAAAGTGGTGAGCAAGGGCATC 2340
Db 2281 CCGGCCCAACAGGGCATCGCGGCAACGAGCAGATCGACAAAGTGGTGAGCAAGGGCATC 2340
Qy 2341 CGCAAGTGCTGTCCTGAGCGGCATCGATGCGCGCATCGTGATCACCAGTACATGAC 2400
Db 2341 CGCAAGTGCTGTCCTGAGCGGCATCGATGCGCGCATCGTGATCACCAGTACATGAC 2400
Qy 2401 GACCTGTACGTGGCAGCGCGGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCAC 2460
Db 2401 GACCTGTACGTGGCAGCGCGGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCAC 2460
Qy 2461 GGTGAATTC 2469
Db 2461 GGTGAATTC 2469

RESULT 2
ABL39960

Id ABL39960 standard; DNA; 2463 BP.

XX ABL39960;
XX
XX
XX

DT 15-MAY-2002 (first entry)
XX

DE Synthetic construct PR975YM SEQ ID NO:31.
XX

XX Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX

OS Human immunodeficiency virus type C.
OS Synthetic.

XX WO200204493-A2.
XX

XX 17-JAN-2002.
XX

XX 05-JUL-2001; 2001WO-US21241.
XX

XX 05-JUL-2000; 2000US-0610313.
XX

XX (CHIR) CHIRON CORP.
PA

PA (UYST-) UNIV STELLENBOSCH.
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX WPI; 2002-154920/20.
XX
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
PT cell lines, particularly in gene therapy -
XX
XX Claim 1; Fig 9; 233pp; English.
XX
XX The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
CC Nef (i). (i) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC AB440054 and ABB06204 to ABB06215 represent sequences used in the
XX exemplification of the present invention.
XX
XX Sequence 2463 BP; 567 A; 835 C; 759 G; 302 T; 0 other;

Query Match 98.9%; Score 2442.2; DB 24; Length 2463;
Best Local Similarity 99.6%; Pred. NO. 1.9e-297;
Matches 2460; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

Qy 1 GTGACGCGCACCATGCGCGAGGCCATGAGCCAGGCCAGCCAGCCCAACTCTCTGATGCGAG 60
Db 1 GTGACGCGCACCATGCGCGAGGCCATGAGCCAGGCCAGCCAGCCCAACTCTCTGATGCGAG 60
Qy 61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGGGCAAGGAGGCG 120
Db 61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGGGCAAGGAGGCG 120
Qy 121 CACATCGCGCCCAACTGCGCGGCCCGCCCGCAGAGAGGGCTGCTGAACTGCGGCAAGGAG 180
Db 121 CACATCGCGCCCAACTGCGCGGCCCGCCCGCAGAGAGGGCTGCTGAACTGCGGCAAGGAG 180
Qy 181 GCGCACCATGAGTGCACCGCGAGCGCGCAGGCGCAACTTCTCCGCGAGGACCTGGCC 240
Db 181 GCGCACCATGAGTGCACCGCGAGCGCGCAGGCGCAACTTCTCCGCGAGGACCTGGCC 240
Qy 241 TTCCCGCGAGGCGCGCGAGTTCGCCAGCGAGAGAACCGCGCCCAACAGCCCGCCACC 300
Db 241 TTCCCGCGAGGCGCGCGAGTTCGCCAGCGAGAGAACCGCGCCCAACAGCCCGCCACC 300
Qy 301 AGCGCGAGCTGCAGGTGCGCGGCGACAAACCCCGCAGGAGGCGCGCGCGCGCGCAG 360
Db 301 AGCGCGAGCTGCAGGTGCGCGGCGACAAACCCCGCAGGAGGCGCGCGCGCGCGCAG 360
Qy 361 GGCACCTTGAACCTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGCTGAGCATCAAGTGTG 420
Db 361 GGCACCTTGAACCTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGCTGAGCATCAAGTGTG 420
Qy 421 GCGCGCCAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGCAGCACCGTGTGGAGGAG 480
Db 421 GCGCGCCAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGCAGCACCGTGTGGAGGAG 480
Qy 481 ATGAGCTTGGCGCGCAAGTGGAGCCCAAGATGATCGGCGCATCGGCGGCTTCTATCAAG 540
Db 481 ATGAGCTTGGCGCGCAAGTGGAGCCCAAGATGATCGGCGCATCGGCGGCTTCTATCAAG 540
Qy 541 GTGCGCCAGTACGACCGAGATCCTGATCGAGATCTCGCGCAAGAGGCCATCGGACCGCTG 600
Db 541 GTGCGCCAGTACGACCGAGATCCTGATCGAGATCTCGCGCAAGAGGCCATCGGACCGCTG 600
Qy 601 CTGATCGGCGCCCGCGCGCGTGAACATCATCGGCGCGCAACATGCTGACCGAGCTGGGTGC 660
Db 601 CTGATCGGCGCCCGCGCGCGTGAACATCATCGGCGCGCAACATGCTGACCGAGCTGGGTGC 660

Qy	661	ACCTTGAACCTTCCCATCAGCCCATCGAGCCATCGAGACCGTGCCTGTAAGCTGAAGCCGCGCATC	720
Db	661		
	661	ACCTTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGTAAGCTGAAGCCGCGCATC	720
Qy	721	GACGGCCCAAGGTGAAGCAGTGGCCCTGTACCGAGGAGAGATCAAGGCCCTGACCGCC	780
Db	721		
	721	GACGGCCCAAGGTGAAGCAGTGGCCCTGTACCGAGGAGAGATCAAGGCCCTGACCGCC	780
Qy	781	ATCTGCGAGGAGATSGAAGAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTAC	840
Db	781		
	781	ATCTGCGAGGAGATSGAAGAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTAC	840
Qy	841	AACACCCCGCTGTTCCCATCAAGAAAGAGACAGCACCAAGTGGCGCAAGCTGGTGGAC	900
Db	841		
	841	AACACCCCGCTGTTCCCATCAAGAAAGAGACAGCACCAAGTGGCGCAAGCTGGTGGAC	900
Qy	901	TTCGCGAGCTGAACAAGCGCACCCAGGACTCTTGGGAGGTGCAAGCTGGGCATCCCCAC	960
Db	901		
	901	TTCGCGAGCTGAACAAGCGCACCCAGGACTCTTGGGAGGTGCAAGCTGGGCATCCCCAC	960
Qy	961	CCGCGCGGCTGAAGAAGAAGACGCTGACCGTGTGGACGTGGCGGACGCTACTTC	1020
Db	961		
	961	CCGCGCGGCTGAAGAAGAAGACGCTGACCGTGTGGACGTGGCGGACGCTACTTC	1020
Qy	1021	AGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCACGATCAAC	1080
Db	1021		
	1021	AGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCACGATCAAC	1080
Qy	1081	AACGAGACCCCGGATCCGCTACCAGTACAAAGTGTGCCCGCAGGCTGGAAAGGCGAGC	1140
Db	1081		
	1081	AACGAGACCCCGGATCCGCTACCAGTACAAAGTGTGCCCGCAGGCTGGAAAGGCGAGC	1140
Qy	1141	CCAGCATCTTCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCGCGGCCCGCAACCCC	1200
Db	1141		
	1141	CCAGCATCTTCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCGCGGCCCGCAACCCC	1200
Qy	1201	GAGATCGTGATCTACCAGTACATGACGACGACCTGTACGTGGGCGACGACCTGGAGATCGGC	1260
Db	1201		
	1201	GAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGGCGACGACCTGGAGATCGGC	1254
Qy	1261	CAGACCGGCGCAAGATCGAGGAGTGGCGCAAGCACCTGCTGCCCTGGGCGTTCAACACC	1320
Db	1255	CAGACCGGCGCAAGATCGAGGAGTGGCGCAAGCACCTGCTGCCCTGGGCGTTCAACACC	1314
	1321	CCGCAAGAAGCACCAAGAAGGAGCGCCCTCTCTGTGGATGGCTACGAGCTGCACCCC	1380
Db	1315	CCGCAAGAAGCACCAAGAAGGAGCGCCCTCTCTGTGGATGGCTACGAGCTGCACCCC	1374
	1381	GACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAAGAGAGCTGGACCGTGAAGCAC	1440
Db	1375	GACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAAGAGAGCTGGACCGTGAAGCAC	1434
	1441	ATCCAGAGCTGTGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTG	1500
Db	1435	ATCCAGAGCTGTGGCGAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGGTG	1494
	1501	CGCCAGCTGTGCAAGCTGTGCGCGCGCGCAAGGCCCTTGACCGACATCTGTCGCCCTGACC	1560
Db	1495	CGCCAGCTGTGCAAGCTGTGCGCGCGCGCAAGGCCCTTGACCGACATCTGTCGCCCTGACC	1554
	1561	GAGAGGCCAGCTGGAGCTGGCCGAGAACCCGAGATCTCTGGCGAGCCCGTGACCGCC	1620
Db	1555	GAGAGGCCAGCTGGAGCTGGCCGAGAACCCGAGATCTCTGGCGAGCCCGTGACCGCC	1614
	1621	GTGTACTACGACCCAGCAAGGACCTGTGTGGCGCGAGATCCAGAAAGAGGGCGCACGACG	1680
Db	1615	GTGTACTACGACCCAGCAAGGACCTGTGTGGCGCGAGATCCAGAAAGAGGGCGCACGACG	1674
	1681	TGGACCTTACCAGATCTACCAGGAGCCCTTTCAAGAACCTTGAAGACCGGCAAGTAGCCCAAG	1740
Db	1675	TGGACCTTACCAGATCTACCAGGAGCCCTTTCAAGAACCTTGAAGACCGGCAAGTAGCCCAAG	1734

XX 05-JUL-2001; 2001WO-US21241.
XX
XX
XX 05-JUL-2000; 2000US-0610313.
XX
XX (CHIR) CHIRON CORP.
XX (UYST-) UNIV STELLENBOSCH.
XX
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX WPI; 2002-154920/20.
XX
XX
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful
XX in applications including DNA immunization or generation of packaging
XX cell lines, particularly in gene therapy -
XX
XX Claim 1; Fig 10; 233pp; English.
XX
XX The present invention describes expression cassettes comprising a
XX polynucleotide sequence encoding a polypeptide comprising immunogenic
XX HIV type C polypeptides. The expression cassettes comprise any of the
XX HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
XX Nef (I). (I) have immunostimulant activity and can be used in gene
XX therapy. The HIV type C polynucleotides are useful in applications
XX including DNA immunisation, generation of packaging cell lines, and
XX production of HIV type C proteins. The polynucleotides are particularly
XX useful in gene therapy and DNA immunisation applications. ABL39942 to
XX ABL40054 and ABL06204 to ABL06215 represent sequences used in the
XX exemplification of the present invention.
XX
XX Sequence 2457 BP; 566 A; 837 C; 754 G; 300 T; 0 other;
XX
XX
XX Query Match 97.8%; Score 2415.4; DB 24; Length 2457;
XX Best Local Similarity 99.3%; Pred. No. 4.2e-294;
XX Matches 2451; Conservative 0; Mismatches 6; Indels 12; Gaps 2;
XX
XX 1 GTCAGGCCACCATGCGCGAGCCATGAGCCAGGCCAGGCCCAACATCTGTATGCGAG 60
XX
XX 1 GTCAGGCCACCATGCGCGAGCCATGAGCCAGGCCAGGCCCAACATCTGTATGCGAG 60
XX
XX 61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGC 120
XX
XX 61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGC 120
XX
XX 121 CACATCGCCCGCAACTGCGCGGCCCGCCCGCAAGAGGGCTGTGAAGTGTGCGCAAGGAG 180
XX
XX 121 CACATCGCCCGCAACTGCGCGGCCCGCCCGCAAGAGGGCTGTGAAGTGTGCGCAAGGAG 180
XX
XX 181 GSCCACCAGATGAAGACTGCAACCGCGCCAGCGCCAGCCCAACTTCTTCGCGAGGACCTGGCC 240
XX
XX 181 GSCCACCAGATGAAGACTGCAACCGCGCCAGCGCCAGCCCAACTTCTTCGCGAGGACCTGGCC 240
XX
XX 241 TTCCCCCGAGGCGCGCGCGAGTTCCTCCAGCGAGCAGAACCGCGCAACGCCCCACCC 300
XX
XX 241 TTCCCCCGAGGCGCGCGCGAGTTCCTCCAGCGAGCAGAACCGCGCAACGCCCCACCC 300
XX
XX 301 AGCCCGGAGCTGCAAGTGTGCGGGCGCAACCCCGCAGCGAGCGCGCGCGAGCGCCAG 360
XX
XX 301 AGCCCGGAGCTGCAAGTGTGCGGGCGCAACCCCGCAGCGAGCGCGCGCGAGCGCCAG 360
XX
XX 361 GSCACCTTGAACCTTCCCGCAGATCCTGTGGACACCGCGCGCGAGCAGCCTGTCTGGAGGAG 420
XX
XX 361 GSCACCTTGAACCTTCCCGCAGATCCTGTGGACACCGCGCGCGAGCAGCCTGTCTGGAGGAG 420
XX
XX 421 GCGGCCAGATCAAGGAGGCGCTGTGTGACACCGCGCGCGAGCAGCCTGTCTGGAGGAG 480
XX
XX 421 GCGGCCAGATCAAGGAGGCGCTGTGTGACACCGCGCGCGAGCAGCCTGTCTGGAGGAG 480
XX
XX 481 ATGAGCCTGCGCGGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGGCGCTTTCATCAAG 540
XX
XX 481 ATGAGCCTGCGCGGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGGCGCTTTCATCAAG 540
XX
XX 541 GTGCCCGAGTACGACCATCTGTATCGAGATCTCGCGCAAGAGGCGCATCGGCGACCGGTG 600

Db 541 GTGCGCCAGTACGACCATCTGATCGAGATCTCGGCAAGAGGCCATCGGCACCGTG 600
QY 601 CTGATCGGCCCAACCCCGTGAACATCATCGGCCGCAACATGTGACCCAGCTGGGCTGC 660
Db 601 CTGATCGGCCCAACCCCGTGAACATCATCGGCCGCAACATGTGACCCAGCTGGGCTGC 660
QY 661 ACCTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGAGCTGAAGCCCGGCATG 720
Db 661 ACCTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGAGCTGAAGCCCGGCATG 720
QY 721 GAGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTTCACCGCC 780
Db 721 GAGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTTCACCGCC 780
QY 781 ATCTGCGAGGATGAGAGAGGAGGCGAGATCACAAGATCGGCCCGGAGAACCCCTAC 840
Db 781 ATCTGCGAGGATGAGAGAGGAGGCGAGATCACAAGATCGGCCCGGAGAACCCCTAC 840
QY 841 AACACCCCGCTGTTCCCATCAAGAAAGAGGAGCAGCAGCAAGTGGCCCAAGCTGGTGGAC 900
Db 841 AACACCCCGCTGTTCCCATCAAGAAAGAGGAGCAGCAGCAAGTGGCCCAAGCTGGTGGAC 900
QY 901 TTCCGCGAGCTGAACAAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGGCATCCGCCAC 960
Db 901 TTCCGCGAGCTGAACAAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGGCATCCGCCAC 960
QY 961 CCGCGCGGCTGAAGAAAGAGAGCGTGCCTGACGCTGTGGACGTGGGCGGACGCTACTTC 1020
Db 961 CCGCGCGGCTGAAGAAAGAGAGCGTGCCTGACGCTGTGGACGTGGGCGGACGCTACTTC 1020
QY 1021 ASCGTGCCCCGTGACGAGGACTTCCGCAAGTACACCGCTTACCATTCCCGCATCAAC 1080
Db 1021 ASCGTGCCCCGTGACGAGGACTTCCGCAAGTACACCGCTTACCATTCCCGCATCAAC 1080
QY 1081 AACGAGACCCCGCGCATCCGTACCACTACAGTGTCTGCCCGAGGCTGGAAGGCGAGC 1140
Db 1081 AACGAGACCCCGCGCATCCGTACCACTACAGTGTCTGCCCGAGGCTGGAAGGCGAGC 1140
QY 1141 CCGAGCATCTCCAGAGCAGCATGACCAAGATCTGCGGCGGCTTCCGCGCGCCCAACCC 1200
Db 1141 CCGAGCATCTCCAGAGCAGCATGACCAAGATCTGCGGCGGCTTCCGCGCGCCCAACCC 1200
QY 1201 GAGATCGTGTATCTACCACTACAGTGTGACGAGCTGTAGTGGGAGCGCTTCCGCGCGCC 1260
Db 1201 GAGATCGTGTATCTACCACTACAGTGTGACGAGCTGTAGTGGGAGCGCTTCCGCGCGCC 1260
QY 1261 GAGACCGCGCGCAAGATCGAGGAGTGTGCAAGCAGCTGTGCGTGGGCGCTTCCAGCAC 1320
Db 1261 GAGACCGCGCGCAAGATCGAGGAGTGTGCAAGCAGCTGTGCGTGGGCGCTTCCAGCAC 1320
QY 1321 CCGGCAAGAACCCAGAGGAGCGCCCTTCCCTGTGGATGGCTACGAGCTGGACCC 1380
Db 1321 CCGGCAAGAACCCAGAGGAGCGCCCTTCCCTGTGGATGGCTACGAGCTGGACCC 1380
QY 1381 GACAAAGTGGACCGTGCAGCCCATCGAGCTGCGCGAGGAGAGCTGAGAGGAGAGAG 1440
Db 1381 GACAAAGTGGACCGTGCAGCCCATCGAGCTGCGCGAGGAGAGAGCTGAGAGGAGAGAG 1440
QY 1441 ATTCAGAGCTGTGGGCAAGCTGAACCTGGGCGAGCCAGATCTACCCCGGCTCAAGGTG 1500
Db 1441 ATTCAGAGCTGTGGGCAAGCTGAACCTGGGCGAGCCAGATCTACCCCGGCTCAAGGTG 1500
QY 1501 CCGCAGCTGTGCAAGCTGTGCGCGCGCCAGAGGCTTACCGACATCTGCGCCCTGACC 1560
Db 1501 CCGCAGCTGTGCAAGCTGTGCGCGCGCCAGAGGCTTACCGACATCTGCGCCCTGACC 1560
QY 1561 GAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGGCGAGCGCTGACGCGC 1620
Db 1561 GAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGGCGAGCGCTGACGCGC 1620
QY 1621 GTGTACTAGACCCCGAGAACCTGTGTGGCGGAGATCCAGAGCGGCGGACAGCCAG 1680
Db 1621 GTGTACTAGACCCCGAGAACCTGTGTGGCGGAGATCCAGAGCGGCGGACAGCCAG 1680

Db 1609 GTGTACTACGACCCCAAGGACCTGGTGGCGGAGATCCAGAGCAGGGCCACGACCAG 1668
 Qy 1681 TGGACCTACAGATCTACCAAGGAGCCCTTCAAGAACTGAAGACCGCAAGTACGCCAAG 1740
 Db 1669 TGGACCTACAGATCTACCAAGGAGCCCTTCAAGAACTGAAGACCGCAAGTACGCCAAG 1728
 Qy 1741 ATGGCACCACCCACCAACCAAGCAGTGAAGCAGCTGACCCAGGCGCTGCGAAGATCGCC 1800
 Db 1729 ATGGCACCACCCACCAACCAAGCAGTGAAGCAGCTGACCCAGGCGCTGCGAAGATCGCC 1788
 Qy 1801 ATGGAGAGATCGTGTATCTGGGGAGAGACCCCAAGTTCCGCTGCCATCCAGAGGAG 1860
 Db 1789 ATGGAGAGATCGTGTATCTGGGGAGAGACCCCAAGTTCCGCTGCCATCCAGAGGAG 1848
 Qy 1861 ACCTGGAGACCTGGTGGACCGGCTACTGCGAGCCACCTGGATCCCGAGTGGGAGTTC 1920
 Db 1849 ACCTGGAGACCTGGTGGACCGGCTACTGCGAGCCACCTGGATCCCGAGTGGGAGTTC 1908
 Qy 1921 GTGAACACCCCCCTGGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGC 1980
 Db 1909 GTGAACACCCCCCTGGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGC 1968
 Qy 1981 GCCGAGACCTTCTACGTGGAGCGGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGC 2040
 Db 1969 GCCGAGACCTTCTACGTGGAGCGGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGC 2028
 Qy 2041 TACGTGACCGACCGGCGCGGAGAGATCGTGAGCTGACCGAGACCAACCAAGCAAG 2100
 Db 2029 TACGTGACCGACCGGCGCGGAGAGATCGTGAGCTGACCGAGACCAACCAAGCAAG 2088
 Qy 2101 ACCGAGCTGACGAGCCATCCAGCTGGCCCTGCGAGACAGCGGACGAGGTGAACATCGTG 2160
 Db 2089 ACCGAGCTGACGAGCCATCCAGCTGGCCCTGCGAGACAGCGGACGAGGTGAACATCGTG 2148
 Qy 2161 ACCGAGACCAAGTACGCCCTGGGCATCATCCAGCCCGAGCCCGACAGAGCGAGCGAG 2220
 Db 2149 ACCGAGACCAAGTACGCCCTGGGCATCATCCAGCCCGAGCCCGACAGAGCGAGCGAG 2208
 Qy 2221 CTGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGACCTGAGCTGGGTG 2280
 Db 2209 CTGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGACCTGAGCTGGGTG 2268
 Qy 2281 CCCGCCACAAAGGCATCGCGGCAAGCAGCAGATCGACAACTGGTGGAGCAAGGGCATC 2340
 Db 2269 CCCGCCACAAAGGCATCGCGGCAAGCAGCAGATCGACAACTGGTGGAGCAAGGGCATC 2328
 Qy 2341 CGCAAGCTGCTTCTTGGAGCGCATCGATGGCGGCATCGTGATCTACCACTACATGAC 2400
 Db 2329 CGCAAGCTGCTTCTTGGAGCGCATCGATGGCGGCATCGTGATCTACCACTACATGAC 2388
 Qy 2401 GACCTGTACGTGGGACGCGGCGCTTAGGATCGATTAAAGCTTCCCGGGGCTTAGCAC 2460
 Db 2389 GACCTGTACGTGGGACGCGGCGCTTAGGATCGATTAAAGCTTCCCGGGGCTTAGCAC 2448
 Qy 2461 GGTGAATTC 2469
 Db 2449 GGTGAATTC 2457

RESULT 4

AAA70481
 ID AAA70481 standard; DNA; 2312 BP.

XX AAA70481;
 AC
 XX
 DT 28-NOV-2000 (first entry)
 XX
 DE HIV FS(-)_ProtMod_Rtopt(+) coding sequence.
 XX
 KW HIV-1; AIDS; Gag-reverse transcriptase; vaccine; expression cassette; ss.
 XX
 OS Human immunodeficiency virus type 1.
 OS Synthetic.

XX WO200039302-A2.
 PN 06-JUL-2000.
 PD
 XX 30-DEC-1999; 99WO-US31245.
 PF
 XX 31-DEC-1998; 98US-0114495.
 PR 01-DEC-1999; 99US-0168471.
 XX (CHIR) CHIRON CORP.
 PA
 XX Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;
 PI Greer C, Selby M, Walker C;
 XX WPI; 2000-452400/39.
 DR
 XX Expression cassettes encoding the human immunodeficiency virus (HIV)
 PT gag-containing polypeptide useful for vaccinating against HIV
 PT Infections and acquired immunodeficiency syndrome (AIDS) -
 XX Claim 7; Fig 75; 391pp; English.
 XX The present sequence is the coding sequence of a synthetic HIV
 CC gag-reverse transcriptase expression cassette, FS(-)_ProtMod_Rtopt(+).
 CC The gag protein of HIV is needed for the assembly of virus-like
 CC particles. In addition, the gag protein is involved in many stages of the
 CC HIV life cycle, including assembly, virion maturation after particle
 CC release and early post-entry steps in viral replication. The expression
 CC cassette may be used for the recombinant expression of HIV
 CC gag-polypeptides which may then be used to vaccinate against HIV
 CC infection and acquired immunodeficiency syndrome (AIDS).
 XX Sequence 2312 BP; 533 A; 750 C; 732 G; 297 T; 0 other;

Query Match 83.1%; Score 2052; DB 21; Length 2312;
 Best Local Similarity 93.6%; Pred. No. 1.le-248;
 Matches 2165; Conservative 0; Mismatches 135; Indels 12; Gaps 2;
 Qy 170 GCGCAAGAGGGCCACAGATGAAGGACTGCACCGAGCGCGCAACCTTCTTCGCGC 229
 Db 1 GCGGCCCGAAGGACACCAATGAAGATTGCATGAGACAGGCTAATTTCTCCGCG 60
 Qy 230 AGGACCTGGCTTCCCGCAGGGCAAGGCCGCGAGTTCCCGCAGCAGACCGCGCCA 289
 Db 61 AGGACCTGGCTTCCCGCAGGGCAAGGCCGCGAGTTTCAGCAGGAGCAGACCGCGCCA 120
 Qy 290 ACAGCCCCACCGCGAGCTGCAGTGCAGGCGCGGCGGCGGCGGCGGCGGCGGCGG 343
 Db 121 ACAGCCCCACCGCGAGCTGCAGTGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 Qy 344 CCGCGCGCGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 397
 Db 181 CCGCGCGCGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 Qy 398 GCCCGCTGGTGAAGCATCAAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 457
 Db 241 GCCCGCTGGTGAAGCATCAAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 Qy 458 CCGACGACACCGCTGCTGGAGGAGATGAGCCCTGCGCGCGCGCGCGCGCGCGCGCGCG 517
 Db 301 CCGACGACACCGCTGCTGGAGGAGATGAGCCCTGCGCGCGCGCGCGCGCGCGCGCGCG 360
 Qy 518 CCGGATCGCGCGGCTTTCATCAAGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 577
 Db 361 CCGGATCGCGCGGCTTTCATCAAGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 Qy 578 GCAAGAGGCGCATCGGACCGCTGCTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 637
 Db 421 GCCAAGAGGCGCATCGGACCGCTGCTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 Qy 638 ACATGCTGACCGAGCTGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 697

Db 481 ACCTGCTGACCCAGATCGGCTGACCCCTGAACCTTCCCATCAGCCCATCGAGACGGTGC 540
QY 698 CGTGAAGCTGAAGCCCGCATGGACGGCCCAAGGTGAAGAGTGGCCCTTGACCCGAGG 757
Db 541 CCGTGAAGCTGAAGCCCGGATGGACGGCCCAAGGTCAAGCACTGGCCCTTGACCCGAGG 600
QY 758 AGAAGATCAAGGCCCTGACCCGCTATCTCGAGGAGATGGAGAGGAGGCAAGATCACCA 817
Db 601 AGAAGATCAAGGCCCTGTTGGAGATCTGCACCGAGATGGAGAGGAGGCAAGATCACCA 660
QY 818 AGATCGGCCCGAGAACCCCTTACACACCCCGGTGTTCCGCATCAAGAGAGAGACAGCA 877
Db 661 AGATCGGCCCGAGAACCCCTTACACACCCCGGTGTTCCGCATCAAGAGAGAGACAGCA 720
QY 878 CCAAGTGGCGCAAGCTGTGTGACTTCCGCGAGCTCAACAGCGCACCCAGGACTTCTGGG 937
Db 721 CCAAGTGGCGCAAGCTGTGTGACTTCCGCGAGCTCAACAGCGCACCCAGGACTTCTGGG 780
QY 938 AGGTGCAGCTGGGCATCCCCACCCCGCGCTTGAAGAAGAGAGCGCTGACCGTGC 997
Db 781 AGGTGCAGCTGGGCATCCCCACCCCGCGCTTGAAGAAGAGAGCGCTGACCGTGC 840
QY 998 TGGAGCTGGGCGACCCCTACTTACAGCTGCCCTGGAGAGAGACTTCCGCAAGTACACG 1057
Db 841 TGGAGCTGGGCGACCCCTACTTACAGCTGCCCTGGAGAGAGACTTCCGCAAGTACACG 900
QY 1058 CTTTCAACATCCCAAGCATCAACAGAGACCCCGCGCATCCGCTACAGTACACAGCTGC 1117
Db 901 CTTTCAACATCCCAAGCATCAACAGAGACCCCGCGCATCCGCTACAGTACACAGCTGC 960
QY 1118 TGCCCCAGGGCTGGAAGGCGACCCCGCATCTTCCAGAGCAGATGACCAAGATCCCTGG 1177
Db 961 TGCCCCAGGGCTGGAAGGCGACCCCGCATCTTCCAGAGCAGATGACCAAGATCCCTGG 1020
QY 1178 AGCCCTTCGCGCGCCGCAACCCCGAGATCTGATCTACAGTACATGAGCAGCTGTACG 1237
Db 1021 AGCCCTTCGCGCGCGGATCAACACCCCGAGATCTGATCTACAGTACATGAGCAGCTGTACG 1080
QY 1238 TGGGAGGAGCTGGAGATCGGCGAGCACCGCGCAAGATCGAGGAGTGGCGAAGCAC 1297
Db 1081 TGGGAGGAGCTGGAGATCGGCGAGCACCGCACCAAGATCGAGGAGTGGCGAAGCAC 1140
QY 1298 TGCTTGGCTGGGGCTTACACACCCCGACAAAGACACCAAGAGAGAGAGAGAGAGAGAG 1357
Db 1141 TGCTTGGCTGGGGCTTACACACCCCGACAAAGAGACACCAAGAGAGAGAGAGAGAGAG 1200
QY 1358 GGATGGGCTAGAGCTGACCCCGACAGTGGACCGTGCAGCCCATCAGCTGCCCGAGA 1417
Db 1201 GGATGGGCTAGAGCTGACCCCGACAGTGGACCGTGCAGCCCATCAGCTGCCCGAGA 1260
QY 1418 AGGAGAGCTGGACCGTGAACACATCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCC 1477
Db 1261 AGGAGAGCTGGACCGTGAACACATCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCC 1320
QY 1478 AGATCTACCCCGGATCAAGTGGCGCAGCTGTGCAAGCTGCTGCGGCGCCCAAGGCGCC 1537
Db 1321 AGATCTACCCCGGATCAAGTGGCGCAGCTGTGCAAGCTGCTGCGGCGCCCAAGGCGCC 1380
QY 1538 TGACCGGATCGTCCCTGACCGAGGAGCGGAGCTGGAGCTGCCCGAGAACCGCGAGA 1597
Db 1381 TGACCGGAGTATCCCTGACCGAGGAGCGGAGCTGGAGCTGGCGAGAACCGCGAGA 1440
QY 1598 TCCTTCGCGAGCGCTGCACGGCTGTACTACGACCCCGACCAAGAGCTGTGTGGCGCGAGA 1657
Db 1441 TCCTTCGAGGAGCGCTGCACGAGTGTACTACGACCCCGACCAAGAGCTGTGTGGCGCGAGA 1500
QY 1658 TCCAGAGCAGGCGCAGACAGTGGAGCTTACAGAGTCTTACAGAGCGCTTCAAGAAC 1717
Db 1501 TCCAGAGCAGGCGCAGGCGCAGTGGAGCTTACAGAGTCTTACAGAGCGCTTCAAGAAC 1560
QY 1718 TGAAGACCGCAAGTACGCAAGATCGCACCGCCCGACACCAAGAGCTGAGAGCTGA 1777
Db 1561 TGAAGACCGCAAGTACGCGCGATGCGCGCGCGCCACACCAAGAGCTGAGAGCTGA 1620

QY 1778 CCGAGGCGCTGCAGAGATTCGCCATGGAGAGCATCGTGTATCTGGGCAAGACCCCAAGT 1837
Db 1621 CCGAGGCGCTGCAGAGATTCGCCATGGAGAGCATCGTGTATCTGGGCAAGATCCCAAGT 1680
QY 1838 TCCGCTGCGCCATCCAGAGGAGACCTGGAGACCTGGTGGACCGACTACTGGCAGGCCA 1897
Db 1681 TCAAGCTGCCATCCAGAGGAGACCTGGAGCGCTGGTGGATGGAGTACTGGCAGGCCA 1740
QY 1898 CTTGGATCCCGAGTGGAGTTCGTGAACACCCCGCCCTGGTGAAGCTGTGTGATCCAGC 1957
Db 1741 CTTGGATCCCGAGTGGAGTTCGTGAACACCCCGCCCTGGTGAAGCTGTGTGATCCAGC 1800
QY 1958 TGGAGAGGAGCCCATCATCGCGCGCGAGACCTTCTACGTGAGCAGCGCGCCCAAGCGG 2017
Db 1801 TGGAGAGGAGCCCATCATCGCGCGCGAGACCTTCTACGTGAGCAGCGCGCGCCCAAGCGG 1860
QY 2018 AGACCAAGATCCGCAAGCGCGCTAGTGAACCGACCGCGCGCGCAAGAGATCGTGAGCC 2077
Db 1861 AGACCAAGCTGGCAAGCGCGCTAGTGAACCGACCGCGCGCGCAAGAGTGGTGAGCA 1920
QY 2078 TGACCGAGACCAACCAAGAGACCGAGCTGCAGGCCATCCAGCTTGGCCCTTGACGAGCA 2137
Db 1921 TCGCGCGACACCAACCAAGAGACCGAGCTGCAGGCCATCCACCTTGGCCCTTGACGAGCA 1980
QY 2138 GCGGAGCGAGGTGAACATCTGTGACCGAGCAGCAGTACGCCCTGGGCATCATCCAGGCC 2197
Db 1981 GCGGAGCGAGGTGAACATCTGTGACCGAGCAGCAGTACGCCCTGGGCATCATCCAGGCC 2040
QY 2198 AGCCCGCAAGAGCGAGCGAGCTGGTGAACCGACATCATCGAGCAGCTGATCAAGAGG 2257
Db 2041 AGCCCGCAAGAGCGAGCGAGCTGGTGAACCGACATCATCGAGCAGCTGATCAAGAGG 2100
QY 2258 AGAAGTGTACCTGAGCTGGTGGTGGCCCAAGGGCATCGCGCGCAACAGCAGATCG 2317
Db 2101 AGAAGTGTACCTGAGCTGGTGGTGGCCCAAGGGCATCGCGCGCAACAGCAGATCG 2160
QY 2318 ACAAGCTGTGAGCAAGGGCATCCGCAAGGTGCTTCTTCTGACCGCATCGATGGCGGA 2377
Db 2161 ACAAGCTGTGAGCGCGCGCATCCGCAAGGTGCTTCTTCTGACCGCATCGATGGCGGA 2220
QY 2378 TCGTATCTACAGTACATGGACGACCTGTAGTGGGAGCGCGCGCCCTAGGATCGATT 2437
Db 2221 TCGTATCTACAGTACATGGACGACCTGTAGTGGGAGCGCGCGCCCTAGGATCGATT 2280
QY 2438 AAAAGCTTCCCGGGCTAGCAGCCGCGTGAATTC 2469
Db 2281 AAAAGCTTCCCGGGCTAGCAGCCGCGTGAATTC 2312

RESULT 5

ABK91616

ID ABK91616 standard; DNA; 9166 BP.

XX ABK91616;

XX AC

XX XX

DT 14-AUG-2002 (first entry)

XX

-DE Modified HIV protein-encoding plasmid DNA #168.

XX

XX

KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;

KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;

KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX

OS Synthetic.

XX

XX

PN WO200232943-A2.

XX

PD 25-APR-2002.

XX

PF 14-AUG-2001; 2001WO-US25721.

XX

PR 14-AUG-2000; 2000US-225097P.

Db 3495 TGGCCAGTACGACAGATCCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGC 3554
QY 602 TGATCGCCCCACCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGTGCA 661
Db 3555 TGGTGGCCCCACCCCGTGAACATCATCGCGCGCAACCTGCTGACCCAGATCGGTGCA 3614
QY 662 CCTGAACCTTCCCATCAGCCCCATCGAGCCCTGAGCCGTGGAAGCTGAAGCCCGGATGG 721
Db 3615 CCTGAACCTTCCCATCAGCCCCATCGAGCCCTGAGCCGTGGAAGCTGAAGCCCGGATGG 3674
QY 722 ACGGCCCCAAGGTGAAGCAGTGGCCCTGACGAGAGAGATCAAGGCCCTGACGCCCA 781
Db 3675 ACGGCCCCAAGGTGAAGCAGTGGCCCTGACGAGAGAGATCAAGGCCCTGAGGAGA 3734
QY 782 TCTGCGAGGAGATGGGAAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTACA 841
Db 3735 TCTGCGAGGATGGGAAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTACA 3794
QY 842 ACACCCCGTGTTCGCATCAAGAAGGAGACACCAAGTGGCGCAAGCTGGTGGACT 901
Db 3795 ACACCCCGTGTTCGCATCAAGAAGGAGACACCAAGTGGCGCAAGCTGGTGGACT 3854
QY 902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGCTGGCGATCCCCACC 961
Db 3855 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGCTGGCGATCCCCACC 3914
QY 962 CCGCCGCTTGAAGAAGAGAGCGGTGACCGTGTGGACGTGGCGAGCGCTACTTCA 1021
Db 3915 CCGCCGCTTGAAGAAGAGCGGTGACCGTGTGGACGTGGCGAGCGCTACTTCA 3974
QY 1022 GCGTCCCTTGACGAGGACTTCCGCAAGTACACCGCTTTCACCATCCCGAGATCAACA 1081
Db 3975 GCGTCCCTTGACGAGGACTTCCGCAAGTACACCGCTTTCACCATCCCGAGATCAACA 4034
QY 1082 ACAGACCCCGCGCATCCGCTACAGTACAAAGTGTGCCCCAGGCTGGAAGGCGAGCC 1141
Db 4035 ACAGACCCCGCGCATCCGCTACAGTACAAAGTGTGCCCCAGGCTGGAAGGCGAGCC 4094
QY 1142 CCAGCATCTTCCAGACGACATGACCAAGATCCTGGAGCCCTTCCCGCAAGCAGACCCCG 1201
Db 4095 CCAGCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCCGCAAGCAGACCCCG 4154
QY 1202 AGATCGTGTATACCATGAGACGACCTGTACGTGGCGGACGACCTGGAGATCGGCC 1261
Db 4155 ACATCGTGTATACCATGAGACGACCTGTACGTGGCGGACGACCTGGAGATCGGCC 4214
QY 1262 AGCACCGCGCAAGATCGAGGAGCTGCGCAAGCACCTGTGCGCTGGGCTTCACCAACC 1321
Db 4215 AGCACCGCGCAAGATCGAGGAGCTGCGCAAGCACCTGTGCGCTGGGCTTCACCAACC 4274
QY 1322 CCAGCAAGACCAAGAGGAGCCCGCTTCTGTGGATGGGCTACGAGCTGACCCCG 1381
Db 4275 CCAGCAAGAGCACCAAGAGGAGCCCGCTTCTGTGGATGGGCTACGAGCTGACCCCG 4334
QY 1382 ACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAGAGAGAGCTGGACCGTGAACGACA 1441
Db 4335 ACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAGAGAGAGCTGGACCGTGAACGACA 4394
QY 1442 TCAGAAAGCTGTGGCAAGCTGAAGTGGCGGACGACAGATCTACCCCGGCATCAAGGTGC 1501
Db 4395 TCAGAAAGCTGTGGCAAGCTGAAGTGGCGGACGACAGATCTACCCCGGCATCAAGGTGC 4454
QY 1502 GCGAGCTGTGAAGCTGTGCGCGGCGCAAGCCCTGACCGACATCTGTCGCCCTGACCG 1561
Db 4455 GCGAGCTGTGAAGCTGTGCGCGGCGCAAGCCCTGACCGAGGTGGTGGCCCTGACCG 4514
QY 1562 AGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGCGGAGCCCGTGCAGGGCG 1621
Db 4515 AGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGAGGAGAGCCGCTGCAGGGCG 4574
QY 1622 TGTACTACGACCCCGACCAAGGAGCTGGTGGCGGAGATCCAGAAAGACCGGACGACCT 1681
Db 4575 TGTACTACGACCCCGACCAAGGAGCTGATCGCGGAGATCCAGAAAGGAGCGGACGACT 4634

QY 1682 GGACCTACCAGATCTACCGAGGCGCTTCAAGAACCTTGAAGACCGCGCAAGTACGCCAAGA 1741
Db 4635 GGACCTACCAGATCTACCGAGGCGCTTCAAGAACCTTGAAGAACCGCGCAAGTACGCCCGCA 4694
QY 1742 TGGCGACCGCCCAACACAGAGCTGAAGCAGCTGACCGAGGCGGTGACAGAAATCGCCA 1801
Db 4695 TGAAGGCGCCCAACACAGAGCTGAAGCAGCTGACCGAGGCGGTGACAGAAATCGCCA 4754
QY 1802 TGAAGAGCATCTGATCTGGGCAAGACCCCAAGTTCCGCCCTGCCCCATCCAGAAGGAGA 1861
Db 4755 CCAGAGCATCTGATCTGGGCAAGACCCCAAGTTCAAGCTGCCCATCCAGAAGGAGA 4814
QY 1862 CTTGGAGACCTCTGGTGGACGACTACTGCGAGCGCACTGATCCCGAGCTGGGAGTTCCG 1921
Db 4815 CTTGGAGGCTCTGGTGGACGACTACTGCGAGCGCACTGATCCCGAGTTGGAGTTCCG 4874
QY 1922 TGAACACCCCGCTTGGTGAAGCTGTGTACCAAGCTTGGAGAAGGAGCCCATCATCGCGC 1981
Db 4875 TGAACACCCCGCTTGGTGAAGCTGTGTACCAAGCTTGGAGAAGGAGCCCATCATCGCGC 4934
QY 1982 CCGAGACCTTCTAGTGGAGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCCGCGCT 2041
Db 4935 CCGAGACCTTCTAGTGGAGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCCGCGCT 4994
QY 2042 ACCTGACCGACCGCGCGCGAGAAAGATCTGTAGCCCTGACCGAGACCAACCAAGCAAGA 2101
Db 4995 ACCTGACCGACCGCGCGCGAGAAAGATCTGTAGCCCTGACCGAGACCAACCAAGCAAGA 5054
QY 2102 CCGAGTGTGAGGCAATCCAGCTGGCCCTGACAGACAGCGCAGCGAGGTGAACATCGTGA 2161
Db 5055 CCGAGTGTGAGGCAATCCAGCTGGCCCTGACAGACAGCGCCTGGAGGTGAACATCGTGA 5114
QY 2162 CCGACAGCGAGTACGCGCTGGGCGATCATCCAGCCCGCCGACAGAGCGAGCGAGC 2221
Db 5115 CCGACAGCGAGTACGCGCTGGGCGATCATCCAGCCCGCCGACAGAGCGAGCGAGC 5174
QY 2222 TGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAAAGGTGTACCTGAGCTGGGTGC 2281
Db 5175 TGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAAAGGTGTACCTGAGCTGGGTGC 5234
QY 2282 CCGCCCAAGAGGCGATCGCGCGCAAGCAGCAGATCGACAAGCTGGTGAGCAAGGGCATCC 2341
Db 5235 CCGCCCAAGAGGCGATCGCGCGCAAGCAGCAGATCGACAAGCTGGTGAGCGCGCATCC 5294
QY 2342 GCAAGTGTCTTCTGGAGCGGATCGATGGCGGCGATCGTGATCTACCACTA 2393
Db 5295 GCAAGTGTCTTCTGGAGCGGATCGATGGCGGCGATCGTGATCTACCACTA 5346

RESULT 7

ABK91617

ID ABK91617 standard; DNA; 9169 BP.

XX ABK91617;

XX ABK91617;

XX 14-AUG-2002 (first entry)

XX Modified HIV protein-encoding plasmid DNA #169.

DE HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

OS Synthetic.

XX WO200232943-A2.

PN 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.

XX 14-AUG-2000; 2000US-225097P.


```
QY 1682 GGACCTACCAAGATCTACAGAGCCCTTCAAGAACCTGAAGACCGCGCAAGTACGCCAAGA 1741
DB 4643 GGACCTACCAAGATCTACAGAGCCCTTCAAGAACCTGAAGACCGCGCAAGTACGCCGCA 4702
QY 1742 TGGCGACCGCCCAACACAGAGCTGAAGCAGCTGACCGAGGCGCGTGCAAGATCGCCA 1801
DB 4703 TGAAGGGCGCCCAACACAGAGCTGAAGCAGCTGACCGAGGCGCGTGCAAGATCGCCA 4762
QY 1802 TGAAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCCCGCTGCCATCCAGAGAGA 1861
DB 4763 CCGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCAAGCTGCCATCCAGAGAGAGA 4822
QY 1862 CTTGGGAGACCTGCTGGACGACTACTTGGCAGGCCACCTGGATCCCGGAGTGGGTTGG 1921
DB 4823 CTTGGGAGGCTGCTGGACGACTACTTGGCAGGCCACCTGGATCCCGGAGTGGGTTGG 4882
QY 1922 TGAACACCCCCCTTGGTGAAGCTGTGTACCAAGCTGGAGAGGAGGCCCATCATCGGG 1981
DB 4883 TGAACACCCCCCTTGGTGAAGCTGTGTACCAAGCTGGAGAGGAGGCCCATCATCGGG 4942
QY 1982 CCGAGACCTTCTAGTGGAGCGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGCT 2041
DB 4943 CCGAGACCTTCTAGTGGAGCGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGCT 5002
QY 2042 ACCTGACCGACCGGGCGCGCAGAAAGATCTGTAGCCCTGACCGAGACCAACCAAGAGA 2101
DB 5003 ACCTGACCGACCGGGCGCGCAGAAAGTGTGTGCCCTTGACCGACCAACCAAGAGA 5062
QY 2102 CCGAGCTGACAGGCCATCCAGCTGGCCCTGCGAGGACGCGGAGAGTGAACATCGTGA 2161
DB 5063 CCGAGCTGACAGGCCATCCAGCTGGCCCTGCGAGGACGCGGCTGGAGTGAACATCGTGA 5122
QY 2162 CCGACACCGAGTACGCCCTGGGCATCATCAGGCCCGCCGACAGAGCGAGCGAGCAGC 2221
DB 5123 CCGACACCGAGTACGCCCTGGGCATCATCAGGCCCGCCGACAGAGCGAGCGAGCAGC 5182
QY 2222 TGGTGAACCAAGATCATCGAGCAGCTGATCAAGAGGAGAGAGTGTACCTGAGCTGGGTGC 2281
DB 5183 TGGTGAACCAAGATCATCGAGCAGCTGATCAAGAGGAGAGTGTACCTGAGCTGGGTGC 5242
QY 2282 CCGCCCAACAGGCGATCGCGCGGCAACGAGCAGATCGACAGCTGGTGAAGAGGCGATCC 2341
DB 5243 CCGCCCAACAGGCGATCGCGCGGCAACGAGCAGTGGACGCGCTGGTGAAGAGGCGGATCC 5302
QY 2342 GCAAGGTGCTTCTCGAGCGCATCGATGGCGCATCGTGTATCAGCA 2393
DB 5303 GCAAGGTGCTTCTCGAGCGCATCGACAGGCGCCAGGAGGAGCAGAGNA 5354

RESULT 8
ABK91611
ID ABK91611 standard; DNA; 9194 BP.
XX
AC ABK91611;
XX
DT 14-AUG-2002 (first entry)
XX
DE Modified HIV protein-encoding plasmid DNA #163.
XX
KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
XX Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX
OS Synthetic.
XX
PN WO200232943-A2.
XX
PD 25-APR-2002.
XX
PF 14-AUG-2001; 2001WO-US25721.
XX
PR 14-AUG-2000; 2000US-225097P.
```

```
PR 14-NOV-2000; 2000US-252115P.
XX 28-MAR-2001; 2001US-279257P.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADRAPARTI B K.
XX
PI Nabel GJ, Huang Y;
XX
XX WPI; 2002-452382/48.
XX
XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection
XX
XX Disclosure; Page 750-753; 794pp; English.
XX
XX The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
XX Sequence 9194 BP; 2232 A; 2671 C; 2524 G; 1767 T; 0 other;
SQ
Query Match 82.1%; Score 2027; DB 24; Length 9194;
Best Local Similarity 91.6%; Pred. No. 1.3e-245;
Matches 2194; Conservative 0; Mismatches 185; Indels 16; Gaps 4;
QY 14 TGGCCGAGGCGCATGAGCCAGG---CCACAGCGCCACATCTCTGATCGAGCGCAGCAACT 70
DB 2967 TGGCCGAGGCGCATGAGCGAGGTGACCAACAGCGCCACCATCATGATCGAGCGCGCAACT 3026
QY 71 TCAAGGGCCCAACAGCGCATCATCAAGTGTCTCAACTCGGCAAGGAGGCGCACATCGCC 130
DB 3027 TCCGACACAGCGCAAGATCGTGAAGTGTCTCAACTCGGCAAGGAGGCGCACATCGCC 3086
QY 131 GCAACTCGCGCGCCCGCCGCAAGAGGGTGTCTGGAAGTGGCGGCAAGAGGCGCACACAGA 190
DB 3087 GCAACTCGCGCGCCCGCCGCAAGAGGGTGTCTGGAAGTGGCGGCAAGAGGCGCACACAGA 3146
QY 191 TGAAGGACTGCACGAGCGAGCGCAGGCAACTTCTTCCGCGAGGAGCTGGCCCTTCCCGCAGG 250
DB 3147 TGAAGGACTGCACGAGCGAGCGCAGGCTTA-TTTTATGGAAGATCTGGCCCTTCCCGCAGG 3205
QY 251 GCAAGGCGCGGAGTTCGCCAGCGAGCAGAAACCGCCCAACAGCGCCACCGCCCGCAGC 310
DB 3206 GGAAGGCGCAGGAATTTCTTCAGAGCAGACACAGAGCCACAGCCCGCCAGAGAGAGC 3265
QY 311 TGCAGGTGCGGG-----CGACAACCCCGCAGCGAGCGCGCGCGAGCGCGCGCGCA 364
DB 3266 TTCAGGTTGGGGAAGAGACAACAACTCCCTCTCAGAAGCAGAGAGCGCATAGACAGAA 3325
QY 365 -----CCCTGAACCTCCCGCAGATCACCTGTGCGAGCGCCCTGTGTGAGCATCAAG 418
DB 3326 CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCGACCCCTCTGTCAATTAAGA 3385
QY 419 TGGCGCGCGCAGATCAAGAGGCGCTTGTGGACACCGCGCGCGCAGACACCTGTGTGGAGG 478
DB 3386 TAGGGGCGCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGCGCAGACACCTGTGTGGAGG 3445
QY 479 AGATGAGCCTCCCGGCAAGTGGAGGCCAGCAGATGATCGCGCGCATCGCGCGCTTCATCA 538
DB 3446 AGATGAACCTCCCGCGCGCTTGGAGGCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 3505
QY 539 AGGTGCGCGCAGTACGACGAGATCTGTGATCGAGATCTGCGGCAAGAGGCCATCGCGCAGC 598
```



```
QY 1679 AGTGGACCTACCATCTACAGAGCCCTTCAAGACCTGAAGACCGCGCAAGTACGCCA 1738
Db 4646 AGTGGACCTACCATCTACAGAGCCCTTCAAGACCTGAAGACCGCGCAAGTACGCCC 4705
QY 1739 AGATGGCCACCGCCACACCAACGACGTGAAGAGCTGACCGAGGCGCGTGCAGAAGATCG 1798
Db 4706 GCATGAAGGGCGCCACACCAACGACGTGAAGAGCTGACCGAGGCGCGTGCAGAAGATCG 4765
QY 1799 CCATGGAGACATCTGTATCTGGGGCAAGACCCCCCAAGTTCCGCGTCCCATCCAGAAGG 1858
Db 4766 CCACCGAGACATCTGTATCTGGGGCAAGACCCCCCAAGTTCAAGCTGCCATCCATCAGAAGG 4825
QY 1859 AGACCTGGGAGACCTGTGTGACCGACCTACTGTCGAGGCGCCCTGTGATCCCGAGTGGGAGT 1918
Db 4826 AGACCTGGGAGGCTGTGTGACCGACCTACTGTCGAGGCGCCCTGTGATCCCGAGTGGGAGT 4885
QY 1919 TCCTGAACACCCCCCTCTGGTGAAGCTGTGTACCACTGGAGAGAGAGCCCATCATCG 1978
Db 4886 TCCTGAACACCCCCCTCTGGTGAAGCTGTGTACCACTGGAGAGAGAGCCCATCATCG 4945
QY 1979 GCGCGAGACCTTCTAGCTGGAGGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCG 2038
Db 4946 GCGCGAGACCTTCTAGCTGGAGGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCG 5005
QY 2039 GCTACGTGACCGACCGCGCGCGCAGAAAGATCGTGAAGCTGACCGAGACCAACCAAGCAGA 2098
Db 5006 GCTACGTGACCGACCGCGCGCGCAGAAAGATCGTGAAGCTGACCGAGACCAACCAAGCAGA 5065
QY 2099 AGACCGAGCTGACAGGCATCCAGCTGGCCCTGCAGGACACGCGGACGAGGTGAACATCG 2158
Db 5066 AGACCGAGCTGACAGGCATCCAGCTGGCCCTGCAGGACACGCGGCTGAGGTGAACATCG 5125
QY 2159 TGACCGACACGACGACCTGGGCTATCCAGGCGCCGACGCGGCGGACGAGGAGCG 2218
Db 5126 TGACCGACACGACGACCTGGGCTATCCAGGCGCCGACGCGGCGGACGAGGAGCG 5185
QY 2219 AGCTGTGTAACACAGATCATCGACGCTGTATCAAGAGGAGAGGTGTACTGTAGCTGGG 2278
Db 5186 AGCTGTGTAACACAGATCATCGACGCTGTATCAAGAGGAGAGGTGTACTGTAGCTGGG 5245
QY 2279 TGCCCCGCCACAAAGGGCATCGCGGCGCAACGAGCAGATCGACAAAGCTGGTGAAGGCGCA 2338
Db 5246 TGCCCCGCCACAAAGGGCATCGCGGCGCAACGAGCAGGTGGACGGCGCTGGTGAAGGCGCA 5305
QY 2339 TCCGCAAGGTGCTGTTCTTGGACGGCATCGATGCGCGCATCTGTATCCAGTA 2393
Db 5306 TCCGCAAGGTGCTGTTCTTGGACGGCATCGATGCGCGCATCTGTATCCAGTA 5360
```

RESULT 10

```
ABK91612
ID ABK91612 standard; DNA; 9194 BP.
```

```
XX AC ABK91612;
```

```
XX DT 14-AUG-2002 (first entry)
```

```
DE Modified HIV protein-encoding plasmid DNA #164.
```

```
XX KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
XX KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
XX KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
```

```
XX OS Synthetic.
```

```
XX PN WO200232943-A2.
```

```
XX PD 25-APR-2002.
```

```
XX PF 14-AUG-2001; 2001WO-US25721.
```

```
XX PR 14-AUG-2000; 2000US-225097P.
```

```
PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADRABARTI B K.
XX Nabel GJ, Huang Y;
XX WPI; 2002-452382/48.
XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection
XX Disclosure; Page 753-756; 794pp; English.
XX The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX SQ Sequence 9194 BP; 2232 A; 2672 C; 2523 G; 1767 T; 0 other;
```

```
Query Match 82.0%; Score 2025.4; DB 24; Length 9194;
Best Local Similarity 91.6%; Pred. No. 2e-245;
Matches 2193; Conservative 0; Mismatches 186; Indels 16; Gaps 4;
```

```
QY 14 TGGCCGAGGCGCATGAGCCAGS---CCACCAGCGCCACATCTCTGATGCGAGCGCAACT 70
Db 2967 TGGCCGAGGCGCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGAGCGCGCAACT 3026
QY 71 TCAGGCGCCCAAGCGCATCATCAAGTGTCTCAACTGGGCAAGAGGGCCACATCGGCC 130
Db 3027 TCCCAACAGCGCAAGATCGTGAAGTGTCTCAACTGGGCAAGAGGGCCACATCGGCC 3086
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGAAGAGGGCCACACAGA 190
Db 3087 GCAACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGAAGAGGGCCACACAGA 3146
QY 191 TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCCTTCCCGCAGG 250
Db 3147 TGAAGGACTGCACCGAGCGCACAGGCTAA-TTTTATTAGGGAAGATCTGGCCCTTCCACAAG 3205
QY 251 GCAAGGCGCCGAGTTCCCGAGCGAGCAGACCGGCCCAACAGCCCGCCAGCGCGGAGC 310
Db 3206 GGAAGGCGCAGGGAATTTTCTTTCAGAGCAGACGAGCCCAACAGCCCGCCAGCAGAGAGC 3265
QY 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGCGCGCGCGCGCGCAGGCGCA 364
Db 3266 TTCAGGTTTGGGGAAGAGACAACTCCCTCTCAGAGAGCAGAGCCGATAGACAGGAA 3325
QY 365 -----CCCTGAAGTTCCTCCAGATCACTCTGTGGCAGCGCCCTGTGTGAGCATCAAGG 418
Db 3326 CTGTATCTTCTAGCTTCCCTCAGATCACTTGTGGCAGCGACCCCTCTGTCACAATAAGA 3385
QY 419 TGGCGGCGCAGATCAAGAGGCGCTTGTGGACACCGCGCGCGCGCGCGCGCGTGTGGAGG 478
Db 3386 TAGGGGCGCAGCTGAAGAGGAGGCGCTTCTAGACACCGCGCGCGCGCGCGCGTGTGGAGG 3445
QY 479 AGATGAGCTGCCCGCAAGTGGAAAGCCCAAGATGATCGCGGCGCATCGCGGCTTCATCA 538
Db 3446 AGATGAACCTGCCCGCGCGCTGGAAAGCCCAAGATGATCGCGGCGCATCGCGGCTTCATCA 3505
QY 539 AGGTGCGCGCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

Db 3506 AGGTGGCCAGTAGACACAGATCCTGATCGAGATCTGCGGCCACAAGCCATCGGCACCG 3565
QY 599 TGTGTGATCGGCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCT 658
Db 3566 TGTGTGCGGCCACCCCGTGAACATCATCGGCCGCAACCTGCTGACCCAGATCGGCT 3625
QY 659 GCACCTGAATTTCCCATCAGCCCCATCGAGACCGTGCCTGTAAAGCTGAAGCCCGGCA 718
Db 3626 GCACCTGAATTTCCCATCAGCCCCATCGAGACCGTGCCTGTAAAGCTGAAGCCCGGCA 3685
QY 719 TGGACGGCCCCAAGCTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 3686 TGGACGGCCCCAAGCTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGTGTG 3745
QY 779 CCATCTGCGAGGAGATGGAGAAGGAGGCAAGATCAACCAAGATCGCGCCGAGAACCCCT 838
Db 3746 AGATCTGCACCGAGATGGAGAAGGAGGCAAGATCAAGAGATCGGCCCGGAGAACCCCT 3805
QY 839 ACAACACCCCGCTTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTG 898
Db 3806 ACAACACCCCGCTTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTG 3865
QY 899 ACTTCCGCGAGCTGAACAGCGCACCCAGAGACTTCTGGGAGTGCAGCTGGGCATCCCC 958
Db 3866 ACTTCCGCGAGCTGAACAGCGCACCCAGAGACTTCTGGGAGTGCAGCTGGGCATCCCC 3925
QY 959 ACCCCGCGGCTTGAAGAAGAAGAGGCTGACCGTGTGGACGTGGCGACGCCCTACT 1018
Db 3926 ACCCCGCGGCTTGAAGAAGAAGAGGCTGACCGTGTGGACGTGGCGACGCCCTACT 3985
QY 1019 TCAGCTGCCCTTGGACGAGGACTTCCGAAGTACACCGCTTCACCATCCCCAGCATCA 1078
Db 3986 TCAGCTGCCCTTGGACGAGGACTTCCGAAGTACACCGCTTCACCATCCCCAGCATCA 4045
QY 1079 ACAACAGACCCCGCATCCGCTACCAAGTACACCTGTGCCCCAGGGTGAAGGGCA 1138
Db 4046 ACAACAGACCCCGCATCCGCTACCAAGTACACCTGTGCCCCAGGGTGAAGGGCA 4105
QY 1139 GCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGTGAGCCCTTCCGCGCCCGCAACC 1198
Db 4106 GCCCAGCATCTTCCAGTGCAGCATGACCAAGATCCTGTGAGCCCTTCCGCAAGCACAACC 4165
QY 1199 CCGAGATCGTGATACCAAGTACATGGACGACCTGTACGTGGGACGCGACCTGGAGATCG 1258
Db 4166 CCGACATCGTGATACCAAGTACATGGACGACCTGTACGTGGGACGCGACCTGGAGATCG 4225
QY 1259 GCCAGCACCGCCAAAGATCGAGGAGCTGGCAAGCAGCCTGCTGCGCTGGGGCTTCAACA 1318
Db 4226 GCCAGCACCGCCAAAGATCGAGGAGCTGGCCAGCAGCCTGCTGCGCTGGGGCTTCAACA 4285
QY 1319 CCCCAGACAAGACCAAGAGGAGCCCTTCTCTGTGGATGGGCTACGAGCTGCACC 1378
Db 4286 CCCCAGACAAGACCAAGAGGAGCCCTTCTCTGTGGATGGGCTACGAGCTGCACC 4345
QY 1379 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGTGAACG 1438
Db 4346 CCGACAAGTGGACCGTGCAGCCCATCGTGTGCCCGAGAAGGAGACAGCTGGACCGTGAACG 4405
QY 1439 ACATCCAGAAGCTGTGGGCAAGCTGAACCTGGGCGCAGCCAGATCTACCCCGGCATCAAGG 1498
Db 4406 ACATCCAGAAGCTGTGGGCAAGCTGAACCTGGGCGCAGCCAGATCTACCGCGCATCAAGG 4465
QY 1499 TGGCCAGCTGTCAAGCTGTCTGCGGGCGCCAAAGCCCTGACCCGACATCTGTGCCCTGCA 1558
Db 4466 TGGCCAGCTGTCAAGCTGTCTGCGGGCGCCAAAGCCCTGACCCGAGTGTGTGCCCTGCA 4525
QY 1559 CCGAGAGGCCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGCAGGCCGTGCAGC 1618
Db 4526 CCGAGAGGCCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGGAAGGAGGCCGTGCAGC 4585
QY 1619 GCGTGTACTACGACCCAGCAAGGAGCTGGTGGCGGAGATCCAGAGAGAGGCCAGGACC 1678
Db 4586 GCGTGTACTACGACCCAGCAAGGAGCTGTATCGCGAGATCCAGAGAGAGGCCAGGGCC 4645

QY 1679 AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGCAAGGTACGCCA 1738
Db 4646 AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGCAAGGTACGCC 4705
QY 1739 AGATGGCGACCGCCCAACACCAAGCAGCTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCG 1798
Db 4706 GCATGAAGGGCGCCCAACACCAAGCAGCTGAAGCAGCTGACCGAGGCCGTGCAGAAGATCG 4765
QY 1799 CCATGAGAGCATTCTGATCTGGGCAAGACCCCAAGTTCCCGCTGCCCATCCAGAAGG 1858
Db 4766 CCACGAGAGCATCTGATCTGGGCAAGACCCCAAGTTCAAGCTGCCATCCAGAAGG 4825
QY 1859 AGACCTGGGAGACCTGCTGGAGCAGCTACTGCGAGGCCACTGATCCCGAGTGGGAGT 1918
Db 4826 AGACCTGGGAGGCTGCTGGAGCAGCTACTGCGAGGCCACTGATCCCGAGTGGGAGT 4885
QY 1919 TCGTGAACACCCCGCTTGGTGAAGCTGTGTACCAAGCTTGGAGAGGAGGCCCATCATCG 1978
Db 4886 TCGTGAACACCCCGCTTGGTGAAGCTGTGTACCAAGCTTGGAGAGGAGGCCCATCATCG 4945
QY 1979 GCCCGAGAGCTTCTACGTGGAGCGGCCGCCCAACCGCGAGACCAAGATCGGCAAGGCCG 2038
Db 4946 GCCCGAGAGCTTCTACGTGGAGCGGCCGCCCAACCGCGAGACCAAGCTGGCAAGGCCG 5005
QY 2039 GCTACCTGACCGACCGGCCGCGCAGAAAGATCCTGAGCCCTGACCGAGACCAACCAAG 2098
Db 5006 GCTACCTGACCGACCGGCCGCGCAGAAAGTGTGTCCCTTGACCCGACACCAACCAAGCA 5065
QY 2099 AGACCGAGCTGACAGGCATCCAGCTGGCCCTGCAGGACAGCGCAGCGAGGTGAACATCG 2158
Db 5066 AGACCGAGCTGACAGGCATCCACCTGGCCCTGCAGGACAGCGCCTGGAGGTGAACATCG 5125
QY 2159 TGACCGACAGCCAGTACGCCCTGGGCATCATTCAGGCCCGACGCCGACAGAGCGAGAGCG 2218
Db 5126 TGACCGACAGCCAGTACGCCCTGGGCATCATTCAGGCCCGACGCCGCTGGTGGCCCGGCA 5185
QY 2219 AGCTGTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAAAGTGTACCTTGAGCTGGG 2278
Db 5186 AGCTGTGTAGCCAGATCATCGAGCAGCTGATCAAGAGGAGAAAGTGTACCTTGAGCTGGG 5245
QY 2279 TGCCCGCCCAAGAGGCATCGCGCGCAACGACAGCAGATCGACAAGCTGGTGAGCAAGGGCA 2338
Db 5246 TGCCCGCCCAAGAGGCATCGCGCGCAACGAGCAGGTGGACGGCCTGGTGGCCCGGCA 5305
QY 2339 TCCGCAAGCTGTGTCTTGGACGGCATCGATGGCGGCATCGTGTATCTACCACTA 2393
Db 5306 TCCGCAAGCTGTGTCTTGGACGGCATCGACAAGGCCCGCAGGAGGAGCAGAGAA 5360

RESULT 11

AAA70479 standard; DNA; 2306 BP.

AC AAA70479;

DT 28-NOV-2000 (first entry)

DE HIV FS(-)_ProMod_RTopt_YM coding sequence.

KW HIV-1; AIDS; Gag-reverse transcriptase; vaccine; expression cassette; ss.

OS Human immunodeficiency virus type 1.

XX Synthetic.

PN WO200039302-A2.

PD 06-JUL-2000.

PF 30-DEC-1999; 99WO-US31245.

XX 31-DEC-1998; 98US-0114495.

PR 01-DEC-1999; 99US-0168471.

XX (CHIR) CHIRON CORP.
 XX Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;
 PI Greer C, Selby M, Walker C;
 XX WPI; 2000-452400/39.
 DR
 XX Expression cassettes encoding the human immunodeficiency virus (HIV)
 PT Gag-containing polypeptide useful for vaccinating against HIV
 PT Infections and acquired immunodeficiency syndrome (AIDS) -
 XX
 PS Claim 7; Fig 73; 391pp; English.
 XX
 CC The present sequence is the coding sequence of a synthetic HIV
 CC Gag-reverse transcriptase expression cassette, PS(-).ProtMod_Rtopt_YM.
 CC The Gag protein of HIV is needed for the assembly of virus-like
 CC particles. In addition, the Gag protein is involved in many stages of the
 CC HIV life cycle, including assembly, virion maturation after particle
 CC release and early post-entry steps in viral replication. The expression
 CC cassette may be used for the recombinant expression of HIV
 CC Gag-polypeptides which may then be used to vaccinate against HIV
 CC infection and acquired immunodeficiency syndrome (AIDS).
 XX
 SQ Sequence 2306 BP; 529 A; 752 C; 730 G; 295 T; 0 other;

Query Match 82.0%; Score 2025.2; DB 21; Length 2306;
 Best Local Similarity 93.3%; Pred. No. 2.6e-245;
 Matches 2156; Conservative 0; Mismatches 138; Indels 18; Gaps 3;

QY 170 GCGGCAAGAGGGCCACCAGATGAAGACTGCACCGAGCGCCAGGCCAACTTCTTCGCG 229
 DB 1 GCGGCGCGAAGGACACCAAAATGAACATTGCCTGAGAGACAGGCTAAATTTCTCGCG 60

QY 230 AGNACTTGGCTTCCCGCCAGGCGAAGCGCGGAGTTCGCCAGGACAGAACCGGCCA 289
 DB 61 AGNACTTGGCTTCTCGAGGGGAAAGCGCGGAGTTTCAGCAGGAGACAGACCGGCCA 120

QY 290 ACAGCCCCACAGCGGAGCTGCAGTGGCGGCGG-----ACAACCCCGCAGCGAGG 343
 DB 121 ACAGCCCCACCGCGGAGCTGCAGTGTGGGCGCGGAGACACAGCTGAGCGAGG 180

QY 344 CCGGCGCGGAGCGCGGAGGCAACCGCTG-----AACTTCCCGCCAGATCACCGTGTGGCAGC 397
 DB 181 CCGGCGCGGAGCGCGGAGGCAACCGTGAAGTCAACTTCCCGCCAGATCACCGTGTGGCAGC 240

QY 398 GCGCCCTGGTGAAGTCAAGTGGCGGCGGAGTCAAGGAGGCGCTGCTGACACCGCGG 457
 DB 241 GCGCCCTGGTGAAGTCAAGTGGCGGCGGAGTCAAGGAGGCGCTGCTGACACCGCGG 300

QY 458 CCGACGACACCGTGTGGAGGAGATGAGCTGCGCGGCAAGTGAAGCCCAAGATGATCG 517
 DB 301 CCGACGACACCGTGTGGAGGAGATGAGCTGCGCGGCAAGTGAAGCCCAAGATGATCG 360

QY 518 GCGGCATCGCGGCTTCATCAAGTGGCGGAGTACGAGTACGACACAGATCCTGATCGAGATCTGCG 577
 DB 361 GCGGGATCGGGGCTTCATCAAGTGGCGGAGTACGACACAGATCCCGCTGGAGATCTGCG 420

QY 578 GCAGAGAGGCGCATCGGACCGTGTGATCGGCCCGCCCGCTGAACATCATCGCGCGCA 637
 DB 421 GCCACAGGCGCATCGGACCGTGTGTTGGCGCCCGCCCGCTGAACATCATCGCGCGCA 480

QY 638 ACATGCTGACCGAGCTGGCTGCACTCCCGTGAAGTTCAGGCGGCGGCGGCGGCGGCGG 697
 DB 481 ACCTGCTGACCGAGTGGCTGCACTCCCGTGAAGTTCAGGCGGCGGCGGCGGCGGCGG 540

QY 698 CCGTGAAGTGAAGCGCGGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 757
 DB 541 CCGTGAAGTGAAGCGCGGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600

QY 758 AGAAGATCAAGGCGCTGACCGGCTGTCGAGGAGATGAGAGGAGGCGGCGGCGGCGGCGG 817
 DB 601 AGAAGATCAAGGCGCTGTCGAGATGTCACCGGAGATGAGAGGAGGCGGCGGCGGCGGCGG 660

QY 818 AGATCGGCGCGGAGAGCCCTTACAAACACCCCGCTGTTGCGCATCAAGAAGAAGACAGCA 877
 DB 661 AGATCGGCGCGGAGAGCCCTTACAAACACCCCGCTGTTGCGCATCAAGAAGAAGACAGCA 720

QY 878 CCAAGTGGCGCAAGCTGTGTGAGCTTCCGCGAGCTGAACAAAGCGCACCCAGAGCTTCTGGG 937
 DB 721 CCAAGTGGCGCAAGCTGTGTGAGCTTCCGCGAGCTGAACAAAGCGCACCCAGAGCTTCTGGG 780

QY 938 AGTTCAGCTGGGCGATCCCGACCCCGCGGCTTGAAGAAGAAGAAGAGCGCTGACCGTGC 997
 DB 781 AGTTCAGCTGGGCGATCCCGACCCCGCGGCTTGAAGAAGAAGAAGAGCGCTGACCGTGC 840

QY 998 TGGACGTGGCGCGCGCTTACTTTCAGCGCTGCCCTGGAGGAGACTTCCGCAAGTACACCG 1057
 DB 841 TGGACGTGGCGCGCGCTTACTTTCAGCGCTGCCCTGGAGGAGACTTCCGCAAGTACACCG 900

QY 1058 CTTTACCATCCCGAGCATCAACAAAGAGAGCCCGCGGATCCGCTACAGTACAGTGC 1117
 DB 901 CTTTACCATCCCGAGCATCAACAAAGAGAGCCCGCGGATCCGCTACAGTACAGTGC 960

QY 1118 TGCCCCAGGCTGGAAGGGGAGCCCGAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1177
 DB 961 TGCCCCAGGCTGGAAGGGGAGCCCGAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1020

QY 1178 AGCCCTTCCGCGCGCGCAACCCCGAGATCGTGATCTTACAGTACATGGACACCTGTACG 1237
 DB 1021 AGCCCTTCCGCGCGCGCAACCCCGAGATCGTGATCTTACCA-----GGCCCCCTGTACG 1074

QY 1238 TGGCAGCGAGCTGGAGATCGGCCAGCAGCCGCCAAGATCGAGAGAGTGGCGCAGCACC 1297
 DB 1075 TGGCAGCGAGCTGGAGATCGGCCAGCAGCCGCCAAGATCGAGAGAGTGGCGCAGCACC 1134

QY 1298 TGCTGGCGTGGGCTTACACACCCCGCACAGAGACACCAAGAGAGGCGCCCTTCTGT 1357
 DB 1135 TGCTGGCGTGGGCTTACACACCCCGCACAGAGAGACCAAGAGAGGCGCCCTTCTGT 1194

QY 1358 GGATGGGTACGAGCTGCACCCCGCACAGTGGAGCGCTGCAGCCCATGAGTGGCCGAGA 1417
 DB 1195 GGATGGGTACGAGCTGCACCCCGCACAGTGGAGCGCTGCAGCCCATGAGTGGCCGAGA 1254

QY 1418 AGGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGAAGTGAAGTGGGCGCAGCC 1477
 DB 1255 AGGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGAAGTGAAGTGGGCGCAGCC 1314

QY 1478 AGATCTACCGCGCATCAAGTGGCGGAGCTGCAAGCTGCTGGCGGCGCAGGCGCC 1537
 DB 1315 AGATCTACCGCGCATCAAGTGGAGAGCTGTGCAAGCTGCTGGCGGCGCAGGCGCC 1374

QY 1538 TGACCGCATCGTCCCGCTGACCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGA 1597
 DB 1375 TGACCGAGGTGATCCCGCTGACCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGA 1434

QY 1598 TCCTGGCGAGCGCTGACCGGCTGTACTACGACCCCGACAGAGACCTGTGTGGCGGAGA 1657
 DB 1435 TCCTGAAGGAGCGCTGACCGAGTGTACTACGACCCCGACAGAGACCTGTGTGGCGGAGA 1494

QY 1658 TCCAGAAGCGGCGCAGCAGCTGGAGCTTACAGATCTACAGAGCGCTTCAAGAAC 1717
 DB 1495 TCCAGAAGCGGCGCAGGCGGAGCTGGAGCTTACAGATCTACAGAGCGCTTCAAGAAC 1554

QY 1718 TGAAGACCGGCAAGTACGCCAAGATGCGCACCCCGCCACACCAAGACAGCTGAAGCAGCTGA 1777
 DB 1555 TGAAGACCGGCAAGTACGCCAAGATGCGCGGCGCGCCACACCAAGACAGCTGAAGCAGCTGA 1614

QY 1778 CCGAGGCGCTGCAAGAATCGCATGGAGAGATCGTGTATCTGGGGCAAGACCCCAAGT 1837
 DB 1615 CCGAGGCGCTGCAAGAATGAGCAGCAGCGAGAGATCGTGTATCTGGGGCAAGATCCCAAGT 1674

QY 1838 TCCGCTTCCCATCCAGAGGAGACCTGGGAGACCTGTGTGGACCGCTACTTGGGAGGCGCA 1897
 DB 1675 TCAAGCTGCCCATCCAGAGGAGACCTGGGAGGCGCTGTGTGATGGAGTGTGGGAGGCGCA 1734

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or human
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX
PS Disclosure: Page 739-741; 794pp; English.
XX
XX The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ Sequence 9189 BP; 2232 A; 2671 C; 2524 G; 1762 T; 0 other;

Query Match 81.8%; Score 2018.6; DB 24; Length 9189;
Best Local Similarity 91.4%; Pred. No. 1.4e-244;
Matches 2190; Conservative 0; Mismatches 184; Indels 21; Gaps 4;

QY 14 TGCCCGAGGCCATGAGCCAGG---CCACCAGCGCCCAACATCCTGATGACGCGCAGCAACT 70
DB TGCCCGAGGCCATGAGCCAGGTTGACCAAGCGCCCAACATGATGACGCGCGCAACT 3026
QY 71 TGAAGGGCCCAAGCGCATCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATCGCCC 130
DB TCCGCAACAGCGCCAGATCGTGAAGTGTCTCAACTGCGGCAAGGAGGGCCACACCGCCC 3086
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGGTGTCTGGAAGTGTGCGCAAGGAGGGCCACACAGA 190
DB GCAACTGCGCGCGCCCGCCGCAAGAGGGTGTCTGGAAGTGTGCGCAAGGAGGGCCACACAGA 3146
QY 191 TGAAGGACTGCAACGAGCGCGCCAGCAACTTCTCCGCGAGGACCTGGCTTCCCGCCAGG 250
DB TGAAGGACTGCAACGAGCGCGCCAGCAACTTCTCCGCGAGGACCTGGCTTCCCGCCAGG 3200
QY 251 GCAAGGCCCGCGAGTTCCTCCAGCGAGCAGAACCGCCCAACAGCCGCCACAGCGCGGAGC 310
DB GCAAGGCCCGCGAGTTCCTCCAGCGAGCAGAACCGCCCAACAGCCGCCACAGCGAGAGC 3260
QY 311 TGCAGTGTGCGCG- - - - -CGACAACCCCGCAGCAGGCGCGCGCGCGCGCGCGCGCA 364
DB TGCAGTGTGCGCGAGAGACAACTCCCTCTCAGAAAGCAGGAGCGCGATAGACAAAGAA 3320
QY 365 - - - - -CCCTGAACTTCCCGCAGATCACCTGTGCGAGCGCCCGCTGTGAGCATCAAG 418
DB CTGTATCTTTCCTTCCCTCAGATCACCTTGTGCGAGCGACCCCTGTGATCAATAAGA 3380
QY 419 TGGCGCGCAGATCAAGAGGCGCCCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGCG 478
DB TAGGGCGCGCAGCTGAAGAGGCGCCCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCG 3440
QY 479 AGATGAGCTTCCCGCAGAGTGAAGCCCAAGATATCGGCGCGATCGGCGCGCTTCATCA 538
DB AGATGAGCTTCCCGCGCGCGCTGGAAGCCCAAGATATCGGCGCGATCGGCGCGCTTCATCA 3500
QY 539 AGGTGCGCGCAGTACGACCAAGATCTCTGATGATCTGCGGCAAGAGGCCATCGGCGACCG 598
DB AGGTGCGCGCAGTACGACCAAGATCTCTGATGATCTGCGGCGCAAGAGGCCATCGGCGACCG 3560
QY 599 TGTGTATCGGCGCGCCCGCGTGAACATCATCGGCGCGCAACATGTGTGAACCGCGCTGGGT 658
DB TGTGTATCGGCGCGCCCGCGTGAACATCATCGGCGCGCAACATGTGTGAACCGCGCTGGGT 3620
QY 659 GCACCTGTAACCTCCCATCAGCCCATCGAGACCGTGTCCGCTGGAAGCTGAAGCCCGGCA 718
DB GCACCTGTAACCTCCCATCAGCCCATCGAGACCGTGTCCGCTGGAAGCTGAAGCCCGGCA 3680

QY 719 TGGACGGCCCCAAGGTGAACAGTGGCCCTGTACCCGAGGAGAGATCAAGGCCCTGACCG 778
DB TGGACGGCCCCAAGGTGAACAGTGGCCCTGTACCCGAGGAGAGATCAAGGCCCTGTTGG 3740
QY 779 CCATCTGCGAGGAGATGGAGAGGAGGCGCAAGATCAACAAAGATCGGCGCGCGAGAACCCCT 838
DB AGATCTGACCCGAGATGGAGAGGAGGCGCAAGATCAACAAAGATCGGCGCGCGAGAACCCCT 3800
QY 839 ACAACACCCCCGTTTGGCCATCAAGAAAGAGGACACCAAGTGGCGCAAGTGGTGG 898
DB ACAACACCCCCGTTTGGCCATCAAGAAAGAGGAGACACCAAGTGGCGCAAGTGGTGG 3860
QY 899 ACTTCGCGAGCTGAACAAGCGACCCAGGACTTCTGGGAGGTGCAGCTGGGATCCGCC 958
DB ACTTCGCGAGCTGAACAAGCGACCCAGGACTTCTGGGAGGTGCAGCTGGGATCCGCC 3920
QY 959 ACCCCGCGGCTTGAAGAAGAGAGCGTGCAGCGTGTGGAGTGGGCGAGCGCTACT 1018
DB ACCCCGCGGCTTGAAGCAGAGAGAGCGTGCAGCGTGTGGAGTGGGCGAGCGCTACT 3980
QY 1019 TCAGCTGCGCCCTGGACGAGGACTTCGCAAGTACACCGCTTTCACCATCCCGAGCATCA 1078
DB TCAGCTGCGCCCTGGACGAGGACTTCGCAAGTACACCGCTTTCACCATCCCGAGCATCA 4040
QY 1079 ACAACGAGACCCCGCATCCGCTACCACTACAACGTGCTGCCCGCAGGCTGGAAGGCA 1138
DB ACAACGAGACCCCGCATCCGCTACCACTACAACGTGCTGCCCGCAGGCTGGAAGGCA 4100
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCCGCGCGCAACC 1198
DB GCGCCAGCATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCGCTTCCCGCAGCAGNACC 4160
QY 1199 CCAGATCTGTATCTTACCAGTACATGGACGACCTGTACGTGGGCGAGCGACTGGAGATCG 1258
DB CCAGATCTGTATCTTACCAGTACATGGACGACCTGTACGTGGGCGAGCGACTGGAGATCG 4220
QY 1259 GCGAGCACCGCCCAAGATCGAGGAGCTGCGCAAGCAGCATGCTGCGCTGGGCTTACCA 1318
DB GCGAGCACCGCCCAAGATCGAGGAGCTGCGCAAGCAGCATGCTGCGCTGGGCTTACCA 4280
QY 1319 CCGCCGACAGAGCAGCAGGAGCGCCCTTCTCTGTGGATGGGCTAGAGCTGCACC 1378
DB CCGCCGACAGAGCAGCAGGAGCGCCCTTCTCTGTGGATGGGCTAGAGCTGCACC 4340
QY 1379 CCGACAAGTGGACCGTGCAGCCCATCGAGTGCCTGCGAGAGGAGAGCTGAGCGTGAACG 1438
DB CCGACAAGTGGACCGTGCAGCCCATCGAGTGCCTGCGAGAGGAGAGCTGAGCGTGAACG 4400
QY 1439 ACATCCAGAACTGGTGGGCAAGCTGAACCTGGGCGCAGCCAGATCTACCCCGCATCAAG 1498
DB ACATCCAGAACTGGTGGGCAAGCTGAACCTGGGCGCAGCCAGATCTACCCCGCATCAAG 4460
QY 1499 TGGCGCAGCTGTCAAGCTGCTGCGCGGCGCCAGGCGCTGACCGCATCGTCCCGCTGA 1558
DB TGGCGCAGCTGTCAAGCTGCTGCGCGGCGCCAGGCGCTGACCGCATCGTCCCGCTGA 4520
QY 1559 CCGAGAGCGCGAGCTGGAGCTGGCGGAGAACCCCGAGATCTTTCGCGGAGCGCGCTGACG 1618
DB CCGAGAGCGCGAGCTGGAGCTGGCGGAGAACCCCGAGATCTTTCGCGGAGCGCGCTGACG 4580
QY 1619 GCGTGTACTAGACCCCGCAGGAGCTTGGTGGCGCAGATCCAGAGGAGCGGCGCAGGAC 1678
DB GCGTGTACTAGACCCCGCAGGAGCTTGGTGGCGCAGATCCAGAGGAGCGGCGCAGGAC 4640
QY 1679 AGTGGACCTACAGATCTTACCAGGAGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCA 1738
DB AGTGGACCTACAGATCTTACCAGGAGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCA 4700
QY 1739 AGATGCGCACCGCCCAACAGCAGCTGAAGCAGCTGACCGAGGCGCTGACGAAGATCG 1798
DB AGATGCGCACCGCCCAACAGCAGCTGAAGCAGCTGACCGAGGCGCTGACGAAGATCG 4760

QY 1799 CCATGGAGAGCATCGTGTGATCTGGGCGAAGACCCCAAGTTCGGCTGCCATCCAGAGG 1858
DB 4761 CCACCGAGAGCATCGTGTGATCTGGGCGAAGACCCCAAGTTCAGCTGCCATCCAGAGG 4820
QY 1859 AGACCTGGGAGACCTGTGTGACCCACTACTGGCAGGCCACCTGTGATCCCGAGTGGGAGT 1918
DB 4821 AGACCTGGGAGGCTGTGTGACCCAGTACTGGCAGGCCACCTGTGATCCCGAGTGGGAGT 4880
QY 1919 TCGTGAACACCCGCCCTGTGTGAAGTGTGTACAGCTGGAGAGGAGGCCATCATCG 1978
DB 4881 TCGTGAACACCCGCCCTGTGTGAAGTGTGTACAGCTGGAGAGGAGGCCATCATCG 4940
QY 1979 GCGCCGAGACCTTCTACGTGTGACCGCGCGCCCAACCGCAGACCAAGATCGCGAAGGCCG 2038
DB 4941 GCGCCGAGACCTTCTACGTGTGACCGCGCGCCCAACCGCAGACCAAGATCGCGAAGGCCG 5000
QY 2039 GCTACGTGACCGACCGCGCGCGCGCAGAGATCGTGAAGCTGACCGAGACCAACACCA 2098
DB 5001 GCTACGTGACCGACCGCGCGCGCGCAGAGTGTGTGCGCCCTGACCGACACCAACACCA 5060
QY 2099 AGACCGAGCTGCAGCCCATCCAGCTGGCCCTGCAGGACAGCGGAGGAGTGAACATCG 2158
DB 5061 AGACCGAGCTGCAGCCCATCCAGCTGGCCCTGCAGGACAGCGGCTGTGAGTGAACATCG 5120
QY 2159 TGACCGACAGCAGTACGCCCTGGCGCATCATCCAGGCGCCAGCCGACAAAGAGCGAGAGCG 2218
DB 5121 TGACCGACAGCAGTACGCCCTGGCGCATCATCCAGGCGCCAGCCGACAAAGAGCGAGAGCG 5180
QY 2219 AGCTGGTGAACCATCATTCGAGCAGCTGTATCAAGAGGAGGAAGTGTACCTGAGCTGGG 2278
DB 5181 AGCTGGTGAACCATCATTCGAGCAGCTGTATCAAGAGGAGGAAGTGTACCTGAGCTGGG 5240
QY 2279 TGCCCGGCCCAAGGGCATCGCGCGCAACGACAGATCGCAAGCTGTGTGAGCAAGGSCA 2338
DB 5241 TGCCCGGCCCAAGGGCATCGCGCGCAACGACAGATCGCAAGCTGTGTGAGCGCCGCA 5300
QY 2339 TCCGCAAGGTGCTGTCTCTGGACGGCATCGATGGCGGCGCATCGTATCTACCAAGTA 2393
DB 5301 TCCGCAAGGTGCTGTCTCTGGACGGCATCGATGGCGGCGCATCGAAGGCCAGGAGGACCGAGAA 5355

RESULT 14

ABK91619
ID ABK91619 standard; DNA; 9167 BP.
AC ABK91619;
XX
DI 14-AUG-2002 (first entry)
XX
DE Modified HIV protein-encoding plasmid DNA #171.
XX
KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX
OS Synthetic.
XX
XX W0200232943-A2.
XX
PD 25-APR-2002.
XX
PF 14-AUG-2001; 2001W0-US25721.
XX
PR 14-AUG-2000; 2000US-225097P.
PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADRABARTI B K.
XX
PI Nabel GJ, Huang Y;
XX
DB WPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection
XX
PS Disclosure; Page 775-778; 794pp; English.
XX
CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ Sequence 9167 BP; 2225 A; 2668 C; 2517 G; 1757 T; 0 other;

Query Match 81.7%; Score 2017; DB 24; Length 9167;
Best Local Similarity 91.4%; Pred. No. 2.3e-244;
Matches 2189; Conservative 0; Mismatches 185; Indels 21; Gaps 4;

QY 14 TGGCGGAGGCCATGAGCCAGG---CCACCAGCGCCCAACATCTGTGATCGAGCGCACT 70
DB 2964 TGGCGGAGGCCATGAGCCAGGTCACCAAGCGCGCCACCATCATGTGAGCGGCGCACT 3023
QY 71 TCAAGGGGCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGCGCCACATCTGCC 130
DB 3024 TCCGCAACCGCGGCAAGATCGTGAAGTCTTCACTGCGGCAAGGAGGCGCCACCGCCC 3083
QY 131 GCAACTGCG 190
DB 3084 GCAACTGCG 3143
QY 191 TGAAGGACTGACCGAGCG 250
DB 3144 TGAAGGACTGACCGAGCG 3197
QY 251 GCAAGGCGCGCGAGTTCGCCAGCGAGACCGCGCGCAACAGCGCGCGCGCGCGCGCGCGCG 310
DB 3198 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGACAGACAGCGCGCGCGCGCGCGCGCG 3257
QY 311 TGCAGTGGCGCGG-----CGACAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 364
DB 3258 TTCAGGTTTGGGGAAGAGACACAACATCCCTCTCAGAAGCAGGAGCGCGCGCGCGCGCG 3317
QY 365 -----CCCTGAACCTTCCCGCAGATCACCTCTGCGAGCGCGCGCGCGCGCGCGCGCG 418
DB 3318 CTGTATCTTTCCTTTCCTCAGATCATCTTTCGAGCGCGCGCGCGCGCGCGCGCGCGCG 3377
QY 419 TGGCGCGCGAGATCAAGGAGCGCTTGTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 478
DB 3378 TAGGGGCGCAGCTGAAGGAGCGCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCG 3437
QY 479 AGATGAGCTGCGCGCGCAAGTGGAGCGCAAGATGATCGGGGCGATCGGGCGCTTCATCA 538
DB 3438 AGATGAACCTGCGCGCGCGTGGAAAGCGCAAGATGATCGGGGCGATCGGGCGCTTCATCA 3497
QY 539 AGGTGCGCGAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGGCGCGCGCGCGCGCG 598
DB 3498 AGGTGCGCGAGTACGACCATCTCTGATCGAGATCTGCGGCGCAAGAGGCGCGCGCGCG 3557
QY 599 TGCTGATCG 658
DB 3558 TGCTGTTGGCG 3617
QY 659 GCACCTTGAACCTTCCCATCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 718
DB 3618 GCACCTTGAACCTTCCCATCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3677

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 11:44:48 : Search time 3394.25 Seconds
(without alignments)
11780.710 Million cell updates/sec

Title: US-09-610-313-30

Perfect score: 2469

Sequence: 1 gtcgacgcccacatggccga.....gggctagcaccggtgaattc 2469

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpi:**

7: em_estro:**

8: em_htc:**

9: gb_estl:**

10: gb_est2:**

11: gb_htc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: gb_gss:**

18: em_gss_hum:**

19: em_gss_inv:**

20: em_gss_pln:**

21: em_gss_vrt:**

22: em_gss_fun:**

23: em_gss_mam:**

24: em_gss_mus:**

25: em_gss_other:**

26: em_gss_pro:**

27: em_gss_rod:**

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90.8	3.7	2598	11 AY103647	AY103647 Zea mays
2	85.6	3.5	951	13 BM321451	BM321451 rockefell
3	80.2	3.2	1132	13 BM320864	BM320864 rockefell
4	79.6	3.2	1165	13 BM320900	BM320900 rockefell
5	79	3.2	867	13 BM321430	BM321430 rockefell
6	76.6	3.1	1550	13 BM321022	BM321022 rockefell

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	75.4	3.1	2299	11 AY106831	AY106831 Zea mays
8	75	3.0	862	13 BM321023	BM321023 rockefell
9	74.4	3.0	3134	11 AY109500	AY109500 Zea mays
10	73.2	3.0	853	13 BM321393	BM321393 rockefell
c 11	72.8	2.9	566	13 BM587428	BM587428 170006873
c 12	72.4	2.9	801	13 BJ375394	BJ375394 BJ375394
13	71.2	2.9	646	13 BM645449	BM645449 170006873
14	71.2	2.9	924	13 BM584191	BM584191 170006872
c 15	71	2.9	925	17 CNS0091P	AL053013 Drosophila
16	70.6	2.9	861	10 BE366696	BE366696 rockefell
17	70	2.8	494	12 BG349122	BG349122 947029804
18	69.6	2.8	546	14 BQ167279	BQ167279 WHE0062_D
19	69.4	2.8	640	10 BE601575	BE601575 HVSMEH009
20	69.4	2.8	926	14 BQ949605	BQ949605 AGENCOURT
21	69.2	2.8	693	12 BG850115	BG850115 1024028A1
22	69	2.8	901	9 AL551163	AL551163 AL551163
23	68.8	2.8	813	13 BI831786	BI831786 603079057
24	68.8	2.8	819	10 BE040830	BE040830 OF12C07_O
25	68.4	2.8	937	13 BM321010	BM321010 rockefell
26	68.4	2.8	1032	10 BE040732	BE040732 OF05G08_O
27	68.4	2.8	1101	17 CNS00397	AL063912 Drosophila
28	68.4	2.8	1333	13 BM320976	BM320976 rockefell
29	68.2	2.8	596	10 AW909955	AW909955 ur78b09_Y
30	68.2	2.8	799	13 BJ361740	BJ361740 BJ361740
31	68	2.8	500	13 BM372120	BM372120 EBR003_SQ
32	68	2.8	530	14 BQ460810	BQ460810 HA05N15r
33	68	2.8	536	14 BQ469029	BQ469029 HM03C13r
34	68	2.8	538	13 BM368580	BM368580 EBR008_SQ
35	68	2.8	540	9 AJ471121	AJ471121 AJ471121
36	68	2.8	550	10 AV912864	AV912864 AV912864
37	68	2.8	566	14 BQ464692	BQ464692 HF02P20r
38	68	2.8	578	13 BM377112	BM377112 EBR005_SQ
39	68	2.8	579	12 BF253640	BF253640 HVSMEF000
40	68	2.8	582	13 BM372871	BM372871 EBR004_SQ
41	68	2.8	588	14 BQ765498	BQ765498 EBR003_SQ
42	68	2.8	600	9 AL508778	AL508778 AL508778
43	68	2.8	625	14 BQ768295	BQ768295 EBR008_SQ
44	68	2.8	641	12 BG908906	BG908906 Tatr1172B
45	68	2.8	656	13 BM370222	BM370222 EBR008_SQ

ALIGNMENTS

RESULT 1
AY103647
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

AY103647
Zea mays
AY103647
AY103647.1
GI:21206725
HTC.
Zea mays.
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2598)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 2598)
Coe,E.C.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
Location/Qualifiers
1..2598
/organism="Zea mays"
/db_xref="MaizeDB:638378"
/db_xref="taxon:4577"
/clone="PCO142084"

linear HTC 25-MAY-2002


```

ACCESSION   BM321451
VERSION     BM321451.1
KEYWORDS    GI:18055857
SOURCE      Mastigamoeba balamuthi.
ORGANISM    Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE   1 (bases 1 to 951)
AUTHORS     Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
            Philippe,H.
TITLE       The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
            Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
JOURNAL     21819461
MEDLINE
COMMENT     Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, NY 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 951 Std Error: 0.00
            POLYA=No.

FEATURES             Location/Qualifiers
     source           1..951
                     /organism="Mastigamoeba balamuthi"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"

BASE COUNT   186 a 321 c 303 g 139 t 2 others
ORIGIN
Query Match      3.5%; Score 85.6; DB 13; Length 951;
Best Local Similarity 46.0%; Pred. No. 9.9e-05;
Matches 323; Conservative 0; Mismatches 376; Indels 3; Gaps 1;

Qy 1734 GCGCAAGATGCGCACCAGCCACACAGACGATGAAGCAGTGCAGCGAGGCGGTGCAGAA 1793
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 219  GCGCAAGATGCGCAGTGTGCTCGTGCACATCTTCTCGACGAGGACACCGCGCGCGC 278
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1794 GATCGCCATGGAGACATCGTGTATCGTGGGCAAGACCCCAAGTTCGCCCTGCCATCCA 1853
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 279  CATCGCGAGCGCGCGTCTCGGTCTTCGCTTCTGAGGCGCAGACCTCCAGGAGTACTG 338
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1854 GAAGAGACCTGGAGACCTGGTGACCGACTACTGGCAGGCCCACTGGATCCCGCAGTG 1913
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 339  GGAGTGCACCTGAAGGCGCTGTGCTCGGCCCTTACCAGGCGCTCAGATCATCGTCGA 398
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1914 GGAGTGTGTGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGTGGAGAGGCCCAT 1973
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 399  CGAGCGCGGTGACCGGACTCTGTATGATCCACAAAGGGGTTCGCGCGGAGGACACCCCAA 458
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1974 CATCGGCGCCGAGACCTTCTACGTGGAGCGCGCCCAACCGCGAGACCAAGATCGGCAA 2033
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 459  GCTGCTGGAGGACGACGAGGCGCTCGAGGAGTTCGCTTCTCAGACGCTGCTCAAGCA 518
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2034 GCGCGGTACTGTACCGACCGCGCGCGCAGAGATCGTGTGACCTGTGACCGGAGACCA 2093
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 519  GGTCCAGAGGAGGAGCGCGCGCTTCTGSCACAAGATCTCTCCCGAGATCCCGCGGTGTAG 578
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2094 CCAGAAGACCGAGTGCAGGCGCATCCAGCTGGCCCTGCAGGACAGCGCCAGCGAGTGAA 2153
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 579  CGAGGAGACGACGACTGGCGTGTATGAGCTGTACAGCTGTACCGCGGACGGAAGCTGCT 638
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2154 CATGCTGACCGACAGCAGTACGCGCTGGGCATCATCCAGGCCCGCCGACCAAGAGGGA 2213
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 639  GTTCCCGCGCTCAACGTCAACGAC---TTCNTCACCAGAGCAGGTTCGACACATCTA 695
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2214 GAGCGAGTGTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAG 2273
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 696  CGGCTGCGCGCACATCGCTCATCGACGCGCATCAAGCGCGGACCGACGCTGATGCTCGCGG 755
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2274 CTGGGTGCGCGCACAGGCGCATCGCGGCAACGACGAGATCGACAGCTGTGTGAGCAA 2333
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 756 CAAGGTGCGCGCTGCTCGCGGCTACGGCGACGTGGGCAAGGGCTGCGCGGCTGCG 815
Qy 2334 GGGCATCCGCAAGGTGCTCTTCTTGACCGCATCGATGCGGCGATCGTGTATTCAGTA 2393
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 816  CGGCCAGGCTGCGCGTCTCATGTCAGCGAGATCGACCCCATCTGGCGCTGCGGCGTC 875
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2394 CATGACGACACCTGTAGTGGCGAGCGCGCGCTTAGGATCGA 2435
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 876  GATGCGCGCTTCGAGGTCAACACGCTCAGGCGGGGCTCGA 917
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
LOCUS      BM320864
DEFINITION rockefeller.0.46 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
            sequence.
ACCESSION  BM320864
VERSION    BM320864.1
KEYWORDS  GI:18055270
SOURCE    Mastigamoeba balamuthi.
ORGANISM  Mastigamoeba balamuthi
            Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE  1 (bases 1 to 1132)
AUTHORS   Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
            Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
            Philippe,H.
TITLE     The analysis of 100 genes supports the grouping of three highly
            divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
            Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
JOURNAL   21819461
MEDLINE
COMMENT   Contact: Muller Miklos
            Laboratory of Biochemical Parasitology
            The Rockefeller University
            1230 York Avenue, New York, NY 10021, USA
            Email: mmuller@rockvax.rockefeller.edu
            Insert Length: 1132 Std Error: 0.00
            POLYA=No.

FEATURES             Location/Qualifiers
     source           1..1132
                     /organism="Mastigamoeba balamuthi"
                     /strain="ATCC 30984"
                     /db_xref="taxon:108607"
                     /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                     /note="syn: Phreatamoeba balamuthi"

BASE COUNT   220 a 413 c 324 g 171 t 4 others
ORIGIN
Query Match      3.2%; Score 80.2; DB 13; Length 1132;
Best Local Similarity 45.3%; Pred. No. 0.00079;
Matches 375; Conservative 0; Mismatches 444; Indels 9; Gaps 2;

Qy 139  CGCGCCCCCGCAAGAAGGGTGTGGAAGTGCAGAGGCGGCGCCAGATGAAGGAC 198
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 79   CGCGCGAGGCGAAGACGAGCTACCGCGCGGCCANCTGGTGATCCAGGACAAGAACAAG 138
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 199  TGCACCGAGCGCGCAGGCCCACTTCTTCGCGAGGACCTTGGCTTCCCGCAGGCGAGGCC 258
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 139  TACACAGCCCAAGTACCGCTTTCGTC-----GTCCGCTTCACCAACAGGACATCGTC 192
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 259  CGCGAGTTCCCGCAGGAGCAGAACCGCGGCCAACAGCCCCACCGCGCGGAGCTGCAGGTG 318
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 193  TGCCAGATCGCTAGCCCAAGATCGACGCGGACCATCTCTCGCGCGCCCTACTTCGCAC 252
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 319  CGCGCGACACACCCCGCAGGAGCGCGCGCGAGCCGCGGCGGACCCCTGAACTTCCCG 378
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 253  GAGTCAACCGCTCGGCGTCAAGCTCGCGCTGACCAACTACGCGCGCCCTACGCGACT 312
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 379  CAGATCAACCTGTGGCAGCGCCCTGTGTGAGCATCAAGTGGGCGGCGCAGATCAAGGAG 438
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 313  GGCTGTGCTGGCGCGCGCGTGTGCTGAAGAAGCTCAACCTCGACTCCAACTACGAGGT 372
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 439  GCCCTGTGGACACCGCGCGCGGACACACCGGTGCTGGAGGAGATGAGCTGCCCCGCAAG 498
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 373 GTCAAGAGGTCAACGGCGAGGACTACACGTCGAGAGCTCGACAGCGGCCCGCGT 432
QY 499 TGAAGCCCAAGATGATCGCGCGCATCGGGGGTTCATCAAGGTGCGGCAGTACGACCAG 558
Db 433 TTCAAGGCCCTGCTCGACGTCGCGCTGGTCCGCAACCTCGACTGGCGCGCGGTTCGCC 492
QY 559 ATCTGATCGAGATCTGCGCAAGAGGCCATCGGCACCGCTGATCGGCCCGCCACCCC 618
Db 493 GCCTCAAGGGCATGTGCGACGCGCGGTCAACGTCCCC---CACAGCGAGACCGCTTC 549
QY 619 GTGAACATCTCGCGCGCAACATCTGACCCAGCTGGGTGCAACCTGAACTTCCCATC 678
Db 550 GTCGGCTTCAACGCGCAAGAGAGCTCAACGCGCGGTCTCCGCAAGTACATCTTC 609
QY 679 AGCCCATCGAGACCTGCGCGGTGAGCTGAAGCCCGGATGACGCGCCCAAGGTGAAG 738
Db 610 GCGGCGACGTCGCGCGTACATGAAGCTCTCAAGGAGCAGGACGCGCGCTTCGAC 669
QY 739 CAGTGGCCCTGACCGGAGAGATCAAGGCCCTGACCGCCATCTCGGAGGATGGAG 798
Db 670 CGCAGTCTCGCGCTACGCCAAGAGGGTGTCAACGCGGACATGCTCGAAGATCTAC 729
QY 799 AAGGAGGCAAGATCAACAGATCGCGCGCGAGAACCCCTTACACACCCCGGTTCGCC 858
Db 730 ACCGAGGCCACAGCAGATCCGCGCGGACCCGACCTTCGTCCCAAGCGCGCTCGAAG 789
QY 859 ATCAAGAGAGAGACACCAAGTGGCGCAAGCTGGTGGAGTTCGCGAGCTGAACAAG 918
Db 790 CCGAGGGCCCAAGCCCAAGCACTGGGCAAGCGAGGCTGACGTACCAAGCGCGAAG 849
QY 919 CGCACCCAGGACTTCTGGGAGGTGACGTGGGATCCCGCCACCGCGCC 966
Db 850 AACCGGTGCCCAAGAGAGGTCCGCTGGGTACCCCGACGCCGCC 897

RESULT 4
BM320900
LOCUS
DEFINITION
rockefeller.0.353 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
sequence.
ACCESSION
BM320900
VERSION
KEYWORDS
SOURCE
Mastigamoeba balamuthi.
Mastigamoeba balamuthi
Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE
1 (bases 1 to 1165)
Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.
The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
21819461
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 1165 Std Error: 0.00
POLYA-No. Location/Qualifiers
1. .1165
/organism="Mastigamoeba balamuthi"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"
221 a 426 c 336 g 178 t 4 others

FEATURES
source
1. .1165
/organism="Mastigamoeba balamuthi"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"
221 a 426 c 336 g 178 t 4 others

Query Match 3.2%; Score 79.6; DB 13; Length 1165;
Best Local Similarity 45.1%; Pred. No. 0.00099;
Matches 428; Conservative 0; Mismatches 506; Indels 16; Gaps 3;
QY 90 CATCAAGTGTCTCAATGTGGGCAAGAGGGCCACATGCCCCCAACTGCGCGCCCCCG 149
Db 46 CGTCAAGAACAAAGCGTACTTCAAGCGCTTCCAGAGCCAGTTCCGTCGCGCGCGGAGG 105
QY 150 CAAGAGGGGTGTGGAAGTGGCGCAAGAGGGCCACCACATGAAGGACTGACACCGAGCG 209
Db 106 CAAGAGCGGACTACCGCGCGCGCCANTGTGTGATCCAGGACAAGAACAGTACACAGCCC 165
QY 210 CCAGGCCCACTTCTCCGCGAGGACCTGGCCCTTCCCGCAGGGAAGCCCGCGAGTTCCC 269
Db 166 CAAGTACCGGTTGTC-----GTCCGCTTACCACAGGAGACATCGTCTGCCAGATCGC 219
QY 270 CAGCGAGCAGAGCGCGCCCAACAGCCCCACAGCGCGGAGCTGCGAGTGGCGCGGACAA 329
Db 220 CTACGCCAAGATCGAGCGGACCACTCTCGCGCGCGCTACTCGCAGGAGTCAACCGG 279
QY 330 CCCCCGAGCGAGCGCGCGCGCGCGCGCCAGGCGCCCTGAACCTTCCCCCAGATCACCT 389
Db 280 CTTTCGGCGCTCAAGTCCGCGCTGACAACTACGCGCGCGCTACGCGACTGCGCTGCTGCT 339
QY 390 GTGCGAGCGCCCTGTGTGAGCATCAAGGTGGCGCGGCGAGTCAAGGAGGCCCTGCTTGA 449
Db 340 GCGCGCGCGTGTGTGAAGAAGCTCAACCTTACTCCAAGTACGAGGGTGTCAAGAAGGT 399
QY 450 CACCGCGCGCGAGCACACCTGTGTGAGGAGATGAGCTGCCCGGCAAGTGGAGGCCCAA 509
Db 400 CAACGCGGAGGACTACAACTCGAGGAGCTCGACGAGCGCGCGCGCTTCAAGGCC-- 457
QY 510 GATGATCGCGGCGATCGCGGCTTCAATCAAGGTGCGCGCTAGCAGACAGATCTGTATCGA 569
Db 458 --TGCTCGAGCTCGGCTGTGTCGACCTCGACTGGCGCGCGCTGTTCGCGCGCTCAA 515
QY 570 GATCTGGCGCAAGAAGGCCATCGGCACCTGTGTGATCGGCCCGCCACCCCGTGAACATCAT 629
Db 516 GGGCATGTGCGAGCGCGGCTCAACGTCCCCCAGCAGAGACCGCTTCGTTCGGCTTCAA 575
QY 630 CGGCGGCAATGCTGACCCAGCTGGGCTGACCTTCACTTCCCATCAGCCCATCGA 689
Db 576 CGGCGCAAGAAGAGGCTCAACGCGCGCTCTCCGCAAGTACATCTTCGCGCGCGCAGT 635
QY 690 GACGCTGCCGCTGAAGTGAAGCCCGCATGGAGCGCGCCCAAGGTGAAGTGAAGCGCCCT 749
Db 636 CGCGGTATCATGAGCTCTCAAGGAGCAGGAGCGCGCGCTTCGACCGCCAGTTC 695
QY 750 GACCGAGGAGAATCAAGGCCCTTGACCGCATCTGCGAGGAGATGAGAGAGGAGCAA 809
Db 696 GCGCTACGCCCAAGGAGGTGTACCGCGGACATGCTCGAGAAGATCTACACCGAGGCCA 755
QY 810 GATCACCAGATGGCGCGCGCGAGAACCCCTACAACACCCCGCTGTTCGCCATCAGAGAA 869
Db 756 -----CAAGCAGATCGCGCGCGCGGACCTTCGTCCCAAGCGCGCTTCGAGAGCCGA 809
QY 870 GGACGACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGA 929
Db 810 GGGCGCCCAAGCCCAAGACTGGGCAAGCGGAGCTGACGTACCAAGCGCGCAGACCCG 869
QY 930 CTTCTGGGAGTGCAGCTGGGCGATCCCCCAGCGCGCGCTTGAAGAAGAGAGAGCGT 989
Db 870 CGTCGCGCAGAAGAAGGTCCGCTGGCTACCCCGACGCGCGCGCGCGCTAAATTC 929
QY 990 GACGCTGTGAGTGGGCGAGCGCTACTTCAGGCTGCCCTGACGAGG 1039
Db 930 CGGTGCTGTACACCGCCATCTCCGCTGCGGCTTCCGCTGTCNGCG 979

RESULT 5
BM321430
LOCUS
DEFINITION
rockefeller.0.1153 Mastigamoeba balamuthi lambda ZAP II Library
EST 03-JAN-2002

Mastigamoeba balamuthi cDNA similar to ribosomal protein S4, mRNA sequence.
 BM321430
 BM321430.1 GI:18055836
 EST.
 Mastigamoeba balamuthi.
 Mastigamoeba balamuthi.
 Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
 1 (bases 1 to 867)
 Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and Philippe,H.

The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

21819461

Contact: Muller Miklos

Laboratory of Biochemical Parasitology

The Rockefeller University

1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu

Insert Length: 867 Std Error: 0.00

POLYA-No.

Location/Qualifiers

source

1. .867

/organism="Mastigamoeba balamuthi"

/strain="ATCC 30984"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/note="syn: Phreatamoeba balamuthi"

BASE COUNT 185 a 298 c 254 g 124 t 6 others

ORIGIN

Query Match 3.2%; Score 79; DB 13; Length 867;

Best Local Similarity 45.7%; Pred. No. 0.0012;

Matches 314; Conservative 0; Mismatches 370; Indels 3; Gaps 1;

QY 58 CAGCGCAGCACTTCAAGGGCCCAAGCGCATCATCAAGTCTTCACTGCGGCAAGGAG 117

DB 131 CCGCACAAGATGCGGAGTGCCTCCGGTATCATCTGTCGCCACACAGATTGAAGTAC 190

QY 118 GGCACATGCGCCGCACTGCGCGCCCGCCCGCAAGAGGCTCTGGAAGTGGCGCAAG 177

DB 191 CGCTGACCGCGCGTGAAGTCACTGATCGATGATGATGATGATGATGATGATGATG 250

QY 178 GAGGGCCACAGATGAGGACTGACCGCGCCCGCCCGCAAGGCTCTTCCGCGGAGGACTG 237

DB 251 AAGTCCCGCACCGCACACGCACTTCCCGCGGGCTTCTATGGACGCTCGTCTCGATCGACAAG 310

QY 238 GCCTTCCCGCAGGCAAGGCGCGGAGTTCCTCCAGCGAGCAGACCGCGCAACAGCCCG 297

DB 311 ACCGACGAGCACTTCCGCTGCTCTACGACACCAAGGCGCTTCCAGGCGCACCGCATC 370

QY 298 ACCAGCGCGAGCTGCAAGT---CGCGCGGACAAACCCCGCAGCGAGCGCGCGCGAG 354

DB 371 AACTCGACCGAGGCAAGTTCAAGCTCGGCAAGTTCGCGCGCTGCACTGCGCAACAAG 430

QY 355 CGCCAGGCGCCCTGAACTTCCCGGATCACCTGTGCGGACCGCCCTCGTGGTAGCATC 414

DB 431 GGCATCCGTAACCTTGGTGAACCGACGCGCGACGATCGCTTACCCCAACCCCGGATC 490

QY 415 AAGTGGCGGCGGACAGTCAAGGAGGCGCTGCTGCGACACCGCGCGGACACCGTGTG 474

DB 491 AAGTCAACGACGCTCAAGATGACCTGCGTGGGCAAGATCATCGATTCGTCAAG 550

QY 475 GAGGATGAGCTTCCCGGCAAGTGAAGCCCAAGATGATGCGGCGCATCGCGGGTTC 534

DB 551 TTCGAGATCGGCAACCTCGTATGATCATCTACCTGCGGACGCAACCTTGGCGGTCGCGGTC 610

QY 535 ATCAAGTGGCGGAGTACGACGATCGTATCGAGATCTCGGCGCAAGAGCCATCGG 594

DB 611 ATTGTGCGCGGAGACGACGAGGGTCTGTCGATGATCATCCACGTCAGGAGCGCGTC 670

QY 595 ACCGTGTGATCGCGCCACCCCGTGAACATCATTCGGCGCGCAACATGCTGACCCAGCTG 654

DB 671 GGCACACAGTTCGCGCGCGCTGACCAACGCTTCGTGATCGGCAAGGGGACCAAGTCC 730

QY 655 GGCTGCACCCCTGAACCTTCCCATCATCGAGCCCGTGGCGGCGTGAAGCTGAAGCCC 714

DB 731 CTCGTACGCTGCCCGCGGCAAGGCGATCAAGAAGTCGATCATCGAGGAGTTCAGCGCG 790

QY 715 GGCATGACGCGCCCAAGGTGAAGCAG 741

DB 791 CGCCACGCGCACAGGACCGAGCAGGAG 817

RESULT 6

LOCUS BM321022 1550 bp mRNA linear EST 03-JAN-2002

DEFINITION rockefeller.0.1192 Mastigamoeba balamuthi lambda ZAP II Library

Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC

3.3.1.1), mRNA sequence.

ACCESSION BM321022 GI:18055428

VERSION BM321022.1

KEYWORDS EST.

SOURCE Mastigamoeba balamuthi.

ORGANISM Mastigamoeba balamuthi.

REFERENCE 1 (bases 1 to 1550)

AUTHORS Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.

Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,

Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and

Philippe,H.

The analysis of 100 genes supports the grouping of three highly

divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

21819461

Contact: Muller Miklos

Laboratory of Biochemical Parasitology

The Rockefeller University

1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu

Insert Length: 1550 Std Error: 0.00

POLYA-No.

Location/Qualifiers

source

1. .1550

/organism="Mastigamoeba balamuthi"

/strain="ATCC 30984"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/note="syn: Phreatamoeba balamuthi"

BASE COUNT 282 a 528 c 495 g 226 t 19 others

ORIGIN

Query Match 3.1%; Score 76.6; DB 13; Length 1550;

Best Local Similarity 45.6%; Pred. No. 0.0032;

Matches 308; Conservative 0; Mismatches 364; Indels 3; Gaps 1;

QY 1761 CGAGTGAAGCAGCTGACCGAGCGCTGCAGAAGATCCCATGGAGACATCGTATCTG 1820

DB 5 CGGACAGGAGCAGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 64

QY 1821 GGGCAAGACCCCAAGTTCCCGCTGCCCATCCAGAAGGAGACCTGGGAGACCT 1880

DB 65 CGCTGGAAGGCGGAGAACCTCCAGGAGTACTGGAGTGCACCTGGAAGGCGCTGTGTT 124

QY 1881 CGACTACTGGAGCGCACCTGGATCCCGAGTGGAGTTCGTGAACACCCCGCCCTCGT 1940

DB 125 GGGCCCTTACCAAGGCGCTCAGATCATCTGCACGACGCGGTGACGCGACTCTGATGAT 184

QY 1941 GAAGCTGTGTACCAAGCTGGAGAGGAGGCCCATCATCGCGCCGAGACCTTCTACGTGA 2000

DB 185 CCACAAGGGGTTCGGCGCGGAGGACACCCCAAGCTGCTGGAGGACGACGAGGCGCTCGA 244

QY 2001 CGGCGCGCGCAACCGGAGACCAAGATCGGAAGCGCGGCTACGTGACCGCGCGCGCG 2060

DB 245 GGAGTGCCTGCCTCAACAACGTGCTCAAGCAGGTCCAGAAGGAGCAGCCCGGCTTCTG 304


```
QY 1386 GTGACCGTGCAGCCATCGAGTGGCCGAGAGGAGAGTGGACCGTGAACGACATCCA 1445
  ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1320 GAAGACCGCGGCGGTGGACATCTCTCAAGATCATGCTGCTCCACGTACATGGTTCGGCT 1379
QY 1446 GAGCTGTGGCAAGTGAACCTGGGCGCCAGCAGATCTACCCGCGATCAAGTGGCCCA 1505
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1380 GTGCCAGCGCTGGACCTGGCCACCTCTCGAGGAGAACTCAAGAGCCCGCTCAAGAGCTG 1439
QY 1506 GCTGTGAAGCTGCTGGCGGCGCAAGGCGCTGACCGACATCTGCGCCCTGACCGAGGA 1565
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1440 CGTGATGGCGTGGCCAGGAGGTGTGACCAACAGCTGGCGGCCACCTCCACAGCG 1499
QY 1566 GCGCGAGCTGGAGCTGCCGAGAACCCGAGATCTCTCGCGAGCCCGTGCACGCGCTGA 1625
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1500 GCGCTTCAGCGAGAAAGCCCTGTCTACCGCATCGACCGCGAGGCGGTGTACGGGTACTA 1559
QY 1626 CTAGGACCCAGCAAGACCTGTGTGGCGGAGATCCAGAAGCAGGCGCCACGACAGTGGAC 1685
  ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1560 CGAGGACCCCTGCAGCGCCCAACTGCGCCCTGTATGAAGAAGATCGGGCCGTGTGTGGA 1619
QY 1686 CTACAGATCTACGAGGAGCCCTCAAGAACCTGAAGACCGGCAAGTACGCCAGATGCG 1745
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1620 CCA-----CGCCCTCCAGCGCGAGCGGAGAGAGAGCGCCGCTCCGTGTCTC 1673
QY 1746 CACGCGCCACCAACAGCTGAAGCAGCTGACCGAGGCGCTGCAGAAGATTCGCCATGGA 1805
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1674 CAAGATCAACAGGTTTCAGGAGGAGCTGCGCGAGGCGCTGCCAGGAGATGGAGGCGC 1733
QY 1806 GAGCATGCTGATCTGGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAGGAGACTG 1865
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1734 CCGCGTGGCTTCGAGACCGCGCGCGCGGCGATCGCCAAAGATCAAGAGAGAGCGCGTC 1793
QY 1866 GGAGACCTGTGGACGACTACTGGCAGGCGCACCTGATCCCGAGTGGGAGTTCGTGAA 1925
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1794 GTACCGCTGTACCGCTTCATCCCGAGGACCTCGGCGCGGTGTACCTGACCGCGAGAA 1853
QY 1926 CACCCCCCCTGTGAGTGTGGTACCAGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1985
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1854 GCTCAAGTCCCGCGCGAGGAGTCAACAAGGTTCCTGCGCCCTCAGCGAGGCAAGCT 1913
QY 1986 GACCTTCTAGTGGACGCGCGCCGCAA 2012
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1914 CATGACCCCATGCTCGACTGCTCTCA 1940
```

```
RESULT 8
BM321023
LOCUS
DEFINITION
rockefeller.0.594 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
3.3.1.1), mRNA sequence.
ACCESSION
BM321023
VERSION
BM321023.1 GI:18055429
KEYWORDS
EST.
SOURCE
Mastigamoeba balamuthi.
ORGANISM
Mastigamoeba balamuthi
Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE
1 (bases 1 to 862)
AUTHORS
Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.
The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
21819461
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockefeller.edu
Insert Length: 862 Std Error: 0.00
POLYA-No.
Location/Qualifiers
```

```
source
1. .862
/organism="Mastigamoeba balamuthi"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/notes="syn: Phreatamoeba balamuthi"
BASE COUNT 175 a 294 c 260 g 126 t 7 others
ORIGIN

Query Match 3.08; Score 75; DB 13; Length 862;
Best Local Similarity 46.2%; Pred. No. 0.0056;
Matches 28; Conservative 0; Mismatches 327; Indels 3; Gaps 1;

QY 1734 CGCCAAGATGGCACCAGCCACACACGACGTGAAGCATGACCGAGGCGTGCAGAA 1793
  ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 220 CGCCAAGATCGCTGGTGGTCTGCTGCAACATCTTCTGACGAGGACACGCGCGCGCGC 279
QY 1794 GATCGCATGAGAGCATCTGATCTGGGCAAGACCCCAAGTTCCGCCCTGCCATCCA 1853
  ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 280 CATCCGCGAGCGGAGTCTCGGTCTTCGCTGGAAGGCGGAGAACCTCCAGGAGTACTG 339
QY 1854 GAAGGAGACCTGGAGACCTGGTGACGACTACTGGCAGGCCACCTGGATCCCGAGTG 1913
  ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 340 GGAGTGACCTTGAAGGCCCTGTGCTTGGGCCCTACCGAGGACCTCAGANCATCGTGA 399
QY 1914 GGAGTTGCTGAACACCCGCCCTGCTGTAAGCTGTGTACCAGCTGGAGAGGAGGCCAT 1973
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 400 CGACGCGGTGACCGACTCTAATGATCCCAAGGGGTTCGCGCGGAGGACACCCCAA 459
QY 1974 CATCGGCGCGAGACCTTCTTACGTGGAGCGCGCCCAACCGCGAGAGACCAAGATCGGCAA 2033
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 460 GCTGCTGGAGGACGACGAGGCGCTCGAGAGGTTCGCTTCAACAACGTGCTCAAGCA 519
QY 2034 GGCGCGCTAGTGACGACCGCGCGCGCGCAGCAAGATCTGAGCCTGACCGAGAGACCA 2093
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 520 GGTCCAGAAGGAGCAGCCCGCGCTTCTGGCACAAGATCTCCCGAGATCCCGGCTGTCAG 579
QY 2094 CCAGAAGACCGAGCTGACGCGCATCCAGCTGGCCCTGCAGGACAGCGCGAGCGAGTGAA 2153
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 580 CGAGGAGACGACGACTGGCGTGATGAGCTGTACCAGCTGCACCGGACGCAAGCTGCT 639
QY 2154 CATGCTGACGACAGCAGTACGCCCTGGGGCATC---ATCCAGGCGCCAGCCGACAAGAG 2210
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 640 GTTCCCGCGCTCAACGCTCAACGACTCTGTCCACCAAGAGCAAGTTTGAACAATCTACGG 699
QY 2211 CGAGGCGAGCTGCTGAACAGATCATCGACGCTGTATCAAGAGGAGAGGTGTACTCT 2270
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 700 CTGCCGCCACTGCTCATCGACGGCATCAAGCGGCGCACCGACGTGATGCTCGCGCGCAA 759
QY 2271 GAGCTGGTGGCCCGCCCAAGGCGCATCGCGCGCAACGAGCAGATCGACAAGCTGGTGAG 2330
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 760 GGTCCGCTGCTCGCGGCTACGGCGAGTGGCAAGGCTGCGCCGAGTCTGCGCGCGG 819
QY 2331 CAAGGGCATCCG 2343
  ||||| ||||
Db 820 CCAGGGCTGCGC 832

RESULT 9
AY109500
LOCUS
DEFINITION
Zea mays CL1506_1 mRNA sequence.
ACCESSION
AY109500
VERSION
AY109500.1 GI:21213244
KEYWORDS
HTC.
SOURCE
Zea mays.
ORGANISM
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 3134)
AUTHORS
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
```



```

Db 2062 AAGATCCACTGCCCCCTCGTGGGAAGAAGAGCTGTCCGCGAGCTCGGCTCTTCGCG 2121
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1864 TGGGAGACCTGGTGGACCGACTACTGGCAGCCACCTGGATCCCGAGTGGGATTCGTG 1923
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2122 TTCTCGTGGCGTGGCGCCAGGCTGTGGCTCCCTCCCATCGCAGACATCTTCGTC 2181
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1924 AACACCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGGGCC 1983
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2182 TGGCCAGCATTTGAAGGAGCTGTGATGGGCGGTACAGCCACGAGCATGTGGAC 2241
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1984 GAGACCTTCTACGTGGAGCGGCCCAACGCGGAGACCAAGATCGGCAAGCCGGGTAC 2043
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2242 GAGATCGGCCACTTCGACCTCTCGTCAAGGCTTACTTCAAGAACG---AGCACCCCAAG 2298
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 2044 GTGACCGACCGCGCGCGAGAAATCGTGAGCCTGACCGAGACCAACCAAGAAC 2103
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2299 TTCCCAACGCGGCCCTCATGACCCAGTACTCGACTCGCTCCGCGCGCCTACATC 2358
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 2104 GAGTGCAGGCCATTCAGCTGGCCCTGCAGACAGCGCGACGAGGTGAACATCGTGACC 2163
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2359 GACGTCGAAGGCGCGCTTGGCCAGTGGAGTACACCGCGCGCAGCTTGTGTATCAAC 2418
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 2164 GACAGCCAGTACGCCCTGGGCGCATCATCCAGGCCAGCCCGACAAAGAGCGAGCGTG 2223
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2419 GGCAGAGAGCGCCATGGAGCGCGCTCGCCATGATCTCGCGCGAAGTGGGATCACGCC 2478
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 2224 GTGAACAGATCATTCGAGCAGCTGATCAAGAAGAGAGAGGTGTAACCTGAGTGGTGCC 2283
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2479 ATGTACAGATTCACAGCGGCTCTCGGAGCAGCGCGGAGGACACACGAGATGCAC 2538
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 2284 GCCACAGGSCATCGGCGGCAACAGCAGATCGACAGCTGGTGAGCAAGGCGATCCGC 2343
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2539 CTCGTGACGCCAACCGAGGAGGACGACATCTCTCCGCGAGAGCTCGACCGGTGG 2598
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 2344 AAGTGTCTTCTGGAGGCGATCGATGGGCGCATCGTGATCTACCACTACATGAGCAC 2403
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2599 CGCGCGAGTACCCGCGACAGGCTCAAGGTGTGTACGTATCGACCAAGGTAAAGCGCGC 2658
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 2404 CTGTACGTGGCGCGCGGCCCTAGGATCG 2434
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2659 CCGAGGGGTGGAAGTACAGCGTTGGGTTCG 2689
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 10
BM321393
LOCUS BM321393 853 bp mRNA linear EST 03-JAN-2002
DEFINITION rockefeller.0.1222 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
sequence.
ACCESSION BM321393
VERSION BM321393.1 GI:18055799
KEYWORDS EST.
SOURCE Mastigamoeba balamuthi.
ORGANISM Mastigamoeba balamuthi
Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.

REFERENCE 1 (bases 1 to 853)
AUTHORS Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
Gordon,P., Durufle,L., Gaasterland,F., Lopez,P., Muller,M. and
Philippe,H.
The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
21819461
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockefeller.edu
Insert Length: 853 Std Error: 0.00
POLYA-No. Location/Qualifiers
source 1. .853

```

/organism="Mastigamoeba balamuthi"
/strain="ATCC 30384"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"
BASE COUNT 156 a 307 c 263 g 125 t 2 others
ORIGIN
Query Match 3.0% Score 73.2 DB 13; Length 853;
Best Local Similarity 46.9% Pred. No. 0.011;
Matches 299; Conservative 0; Mismatches 329; Indels 9; Gaps 2;
Qy 388 CTGTGGCAGCGCCCTCGTGAGCATCAAGGTGGCGGCCAGATCAAGAGAGCCCTGCTG 447
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 125 CTGCGCCCGCTGTGCTGAAGAAGCTCAACTCGACTCCAAGTAGGAGGTGTCAAGAG 184
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 448 GACACCGCGCGCGACGACACCGTGTGCTGGAGAGATGAGCCTGCCGCGCAAGTGGAGGCC 507
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 185 GTCAACGCGGAGGACTACACGCTGAGAGCTCGACGACGGCGCCGCCGCTTCAAGGCC 244
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 508 AAGATGATCGCGGCATCGCGGCGCTTCATCAAGTGGCGGACGACGACCATCTGTATC 567
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 245 CTGCTCAGCTGGCGCTGTCCGACCTCGACCGCGCGCGCTTTCGCGCCCTCAAG 304
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 568 GAGATCTCGGCAAGAAGGCCATCGGCACCGTGTGATCGCGCCCGCCCGCTGAACATC 627
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 305 GGATGTGCGACGCGCGGCTCAACGTCC---CACACGAGACCGCTTCTCGTGGCTTC 361
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 628 ATCGGCGCACATCTGTACCCAGTGGGTGCGACCTTGAACCTTCCCATCAGCCCCATC 687
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 362 AACGCGCAACAAGAAGAGCTCAAGCGCGGTCTCCGCAAGTACATCTTCGGCGGCCAC 421
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 688 GAGACCGTGGCGCGTGAAGCTGAAGCCCGCATGGACGCGCCCAAGGTGAAGAGGCGCC 747
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 422 GTCCGCGTACATGAAGCTCTCAAGGAGCAGGACGCGCGCTTCGACCGCCAGTTTC 481
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 748 CTGACCGAGGAGAAGATCAAGCGCCCTGACCGCCATCTCGAGAGATGAGAAGAGGCGC 807
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 482 TCAGCTACGCCAAGGAGGTGTACCCCGCACATGCTCGAGAGATCTACACGAGGCC 541
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 808 AAGATCAACAAGATCGGCGCCGAGAACCCCTTACAAACACCCCGTTCGCCATCAAGAG 867
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 542 CA-----CAAGCAGATCCGCGCGACCCGACCTTCGTCGCCAAGCGCGCTCGAAGGCC 595
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 868 AAGCAGACCAAGTGGCGAAGCTGTGTGATCTCCGCGAGCTGAACAGGCGACCCAG 927
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 596 GAGGCGCGCAAGCCCAAGCTGCGGCAAGCGCAGGCTGACGTACCGAGCGCGAAGAAC 655
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 928 GACTTCTGGGAGGTGCAAGTGGGATCCCGACCCCGCGGCTGAAGAAGAAGAGAGC 987
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 656 CGGTGCGCCAGAGAAGGTCCGCTGGGCTACCCCGACGCCCGCCAGCAGTAAATT 715
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 988 GTGACCGTGTGGAGCTGGGCGACGCCCTTACTTCAGCG 1024
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 716 CCGGTCGCCGTATACACNGCCATCTCCCGCTCGCGCG 752
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

RESULT 11
BM587428/c
LOCUS BM587428 566 bp mRNA linear EST 25-FEB-2002
DEFINITION 17000687321202 A.Gam.ad.cDNA blood1 Anopheles gambiae clone
19600449696110 5', mRNA sequence.
ACCESSION BM587428
VERSION BM587428.1 GI:18883289
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoides;
Anopheles.
REFERENCE 1 (bases 1 to 566)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
R., Collins,F.H., Venter,J.C. and Hoffman,S.L.


```

Db 141 ATCATGATCAAGCTCAAGCAGCTCATCAAGTTOAGCTCAAGCAGTTCAGTCAAGCAG 82
QY 1228 GACCTGTACGTGGG 1241
Db 81 TTCTAGTTCAAGGG 68

RESULT 13
BM645449 646 bp mRNA linear EST 26-FEB-2002
LOCUS 17000687316588 A.Gam.ad.cdna1 Anopheles gambiae cdna clone
DEFINITION 19600449648747 5', mRNA sequence.
ACCESSION BM645449
VERSION BM645449.1 GI:18944960
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
REFERENCE 1 (bases 1 to 646)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished (2002)
COMMENT Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU01004920 row: B column: 09
Seq primer: M13 Reverse.
FEATURES
source
1. .646
/organism="Anopheles gambiae"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449648747"
/dev_stage="Adult"
/lab_host="DH10b"
/notes="Vector: pSport1; Site_1: Sall; Site_2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen.
cdna inserts >300 bp cloned directionally into pSport 1.
Not 1 site is 3'. Clones available through the Malaria
Research and Reference Reagent Resource Center
(www.malaria.mr4.org)."
BASE COUNT 170 a 194 c 210 g 72 t
ORIGIN
Query Match 2.9%; Score 71.2; DB 13; Length 646;
Best Local Similarity 48.1%; Pred. No. 0.023;
Matches 202; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

QY 1956 GCTGGAGAGAGCGCCATCATCGCGCGGACCTTCTACGTGGACGGCGCGCCCAACCG 2015
Db 49 GTTCGAGATCGACGCCAACCGCATCTCTGAGGTGTGCGCGGAGGACAAAGGCGCGGCAA 108
QY 2016 CGAGACCAAGATCGGAAGCGCGGTACGTGACGACCGCGGCGCGGAGAGATCGTGAG 2075
Db 109 CCGGAGAGATGCTCATCACCACGACCAAGACCGGTGACGCGCGGACGACATCGAGCG 168
QY 2076 CCTGACCGAGACCAACACAGAGACCGAGCTGCGAGCCATCCAGCTGGCCCTGCAGGA 2135
Db 169 CATGATCAAGGATCGGAGCGGTTTCGCGGACGACGACAAAGAGTGAAGGAGCGCGTGA 228
QY 2136 CAGCGGACGCGGTGACATCTGTCACGACGACCGACGATACGCCCTGGGCATCATCCAGG 2195
Db 229 GGCCCGCAACGAGCTGAGAGCTACGCGTACAGCTGAAGAACAGCTGACGTGCAAGGA 288

```

```

QY 2196 CCAGCCCAAGAGCGAGAGCGAGCTGGTGAACCAAGATCATCGAGCAGCTGATCAAGAA 2255
Db 289 CAAGTGGCGCGGAGCTGTCCGACGACGACAAAGGCCAAGATGGAGGAGCGATCGACGA 348
QY 2256 GGAGAAGTGTACCTGAGCTGGTCCCGCCCAAGAGGCATCGCGCGGACGAGCAGAT 2315
Db 349 GAAGATCAAGTGGCTGGAGACGACGACCGGCGGAGAGTACAAGAGACGAGAA 408
QY 2316 CGACAAGCTGGTGAAGGCGCATCCGCAAGGTGCTGTCTCTGGACGCGCATCGATGGCGG 2375
Db 409 GAAGAGCTGGAGGACATCTGTCACCCCATCATTTCCCAAGCTGTACGCGACGAGTGGCGG 468

RESULT 14
BM584191 724 bp mRNA linear EST 22-FEB-2002
LOCUS 17000687279987 A.Gam.ad.cdna.blood1 Anopheles gambiae cdna clone
DEFINITION 19600449683331 5', mRNA sequence.
ACCESSION BM584191
VERSION BM584191.1 GI:18872658
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
REFERENCE 1 (bases 1 to 724)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished (2002)
COMMENT Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU01004B2V row: B column: 21
Seq primer: M13 Reverse.
FEATURES
source
1. .724
/organism="Anopheles gambiae"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449683331"
/dev_stage="Adult"
/lab_host="DH10b"
/notes="Vector: pSport1; Site_1: Sall; Site_2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cdna inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)."
BASE COUNT 183 a 218 c 243 g 80 t
ORIGIN
Query Match 2.9%; Score 71.2; DB 13; Length 724;
Best Local Similarity 48.1%; Pred. No. 0.024;
Matches 202; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

QY 1956 GCTGGAGAGAGCGCCATCATCGCGCGGACCTTCTACGTGGAGCGCGCGCCCAACCG 2015
Db 224 GTTCGAGATCGACGCCAACCGCATCTCTGAGGTGTGCGCGGAGGACAAAGGCGCGGCAA 283
QY 2016 CGAGACCAAGATCGGAAGCGCGGTACGTGACGACCGCGGCGCGGAGAGATCGTGAG 2075
Db 284 CCGGAGAGATGCTCATCACCACGACCAAGACCGCTGACGCGCGGAGACATCGAGCG 343
QY 2076 CCTGACCGAGACCAACACGAGAGACCGAGCTGCGAGCCATCCAGCTGGCCCTGCAGGA 2135
Db 344 CATGATCAAGGATCGGAGCGGTTTCGCGGACGACGACCAAGAGCTGAAGGAGCGCGTGA 403

```



```
QY 2136 CAGCGGACGAGGTGAACATCTGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGC 2195
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 404 GGCCCGCAACGAGCTCAGAGCTACGGGTACAGCTGAAGAACAGCTCAGCTCGAAGGA 463
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2196 CGAGCCCGCAAGAGGAGGAGCTGTGTGACACAGATCATCGAGCAGCTGATCAAGAA 2255
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 464 CAAGCTGGGCGGAGCGGTGTCCGACGACGACAAAGGCAAGTGAAGGAGCGATCGACGA 523
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2256 GGAGAAGGTGTACTGTAGCTGGTGCCGCCACAGAGGCGCATCGGCGGCAACGAGCAGAT 2315
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 524 GAGATCAAGTGTGCTGACAGAACAGACGACACGAGCGCGGAGAGTACAAAGACAGAA 593
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2316 CGACAAGCTGGTGAAGAGGCGATCCGCAAGGTGCTTCTCGACGCGCATCGATGCGCG 2375
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 584 GAAGGAGCTGGAGACATCTGCAGCCCATCATTTGCCAAGCTGTACGCGAGCAGTGGCGG 643

RESULT 15
CNS0091P/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RPI-98 library from Drosophila melanogaster (fruit
fly); genomic survey sequence.
ACCESSION
AL053013
VERSION
AL053013.1 GI:4934461
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster.
ORGANISM
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 925)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
FEATURES
    source
        1..925
            /organism="Drosophila melanogaster"
            /db_xref="taxon:7227"
            /clone="BACR19D16"
            /clone_lib="RPI-98"
            /note="end : TET3"
BASE COUNT
    120 a      61 c      61 g      172 t      511 others
ORIGIN
Query Match          2.9%; Score 71; DB 17; Length 925;
Best Local Similarity 15.6%; Pred. No. 0.026;
Matches 58; Conservative 173; Mismatches 140; Indels 0; Gaps 0;
QY 141 CGCCCGCCGCAAGAGGCTGCTGAAGTCCGCAAGGAGGCGCACAGATGAAGGACTG 200
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 924 SBSCSCSCSBSBSSSNMTSSNSBSCSSBSBSSTSSMSSSSSGSSSSS 865
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 201 CACCGAGCGCGCAGCAACTTCTTCGCGAGGACCTGGCCTTCCCCCAGGCGGCCG 260
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Db 864 SGTSSACVKNASSSSCGCCCGCMABCCMCMSSSSSCSSASARGVKVRASGAGKRGGS 805
QY 261 CGAGTTCCCGCAGGAGCAGAAACCGCGCCCAACAGCCCCCAGCGAGCTGCAGGTGCG 320
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 804 GASNHSSSSACBSSSSSCSASWSASSSSSRFRSGGAGGSGASSRSSSSSSSA 745
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 321 CGCGCAACACCCCGCAGCAGGCGCGCGCGAGCGCCGAGCGCCACCTGAACCTTCCCCCA 380
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 744 SAGSVSSASSSSSSSSSVSCSVASSMSCSBSSSSSASASSSSSSSASCSCCCT 685
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 381 GATCACCTGTGCGAGCGCCCTCTGTTGAGCATCAAGTGGCGCGCCAGATCAAGGAGGC 440
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 684 SWSCSSTSSMSAARSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSGSACGBS 625
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 441 CCTGCTGGACACCGCGCGCGACACACCTGCTGTGAGGAGATGAGCCTGCCCGCAAGTG 500
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 624 MSSGGSGSVASSGMSVSSSGRSGSGGGGSGSSSSSSSSSGSGSGSVCSVCS 565
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 501 GAAGCCCCAAGA 511
      | : | : |
Db 564 GCMCRCSGSA 554
```

Search completed: February 10, 2003, 20:43:27
Job time : 3412.25 secs

APPLICANT: de BETHUNE, Marie-Pierre
APPLICANT: HERTOGS, Kurt
APPLICANT: PAUWELS, Rudi
TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
FILE REFERENCE: 1377-125P
CURRENT APPLICATION NUMBER: US/09/117,217
CURRENT FILING DATE: 1998-07-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 2601
TYPE: DNA
ORGANISM: HIV-HXB2
FEATURE:
NAME/KEY: CDS
LOCATION: (334)...(489)
OTHER INFORMATION: gag P6 (52 AA)
US-09-117-217-9

Query Match 45.88; Score 1132; DB 4; Length 2601;
Best Local Similarity 58.58; Pred. No. 8.6e-178;
Matches 1626; Conservative 0; Mismatches 730; Indels 16; Gaps 4;
QY 14 TGGCCGAGGCGCATGAGCCAGGCCACCA---GCGCCAAACATCCTGTGCGAGCGCAGCAACT 70
DB 77 TGGCTGAAGCAATGAGCCAGTACAAATTCAGCTACCAATGATGCGAGAGGCGCAAT 136
QY 71 TCAAGGCGCCGAAAGCGCATCAAGTGTCTTCAACTGCGGCAAGAGGCGCCATCGCCC 130
DB 137 TTAGGAACCAAGAAAGATTTGTTAAGTGTTCATTTGTCGAAAGTGTGGAAGGAGGACACCA 196
QY 131 GCAACTGCGCGCCCGCCGGAAGAGGCTGCTGGAAGTGGCGGCAAGGAGGCGCCACAGA 190
DB 197 GAAATTCAGGCGCCCTAGGAAAAGGCTGTTGGAATGTGGAAGGAGGAGGACACCAAA 256
QY 191 TGAAGGACTGCACGAGCGCGCAGGCCCAACTTCTTCCGCGAGGACTGGCCCTTCCCCAGG 250
DB 257 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTGTGGGAGATCTGGCCCTTCTACAG 315
QY 251 GCAAGGCGCGGATTTCCCGACGAGGAGAGAACCGCGCCACAGCCCGCAGCGCGAGC 310
DB 316 GGAAGGCGCGGGAATTTTCTTACAGAGAGACGAGCGCAACAGCGCCCGCAGAGAGAGC 375
QY 311 TGCAGGTGCGGG-----CGACAACCCCGCAGCGGCGGCGCGAGCGCGCAGGGCA 364
DB 376 TTCAGGTCTGGGTGAGAGACAACAACTCCCCCTCAGAGGAGGAGCGGATAGACAAGAA 435
QY 365 -----CCCTGAATCTCCCGAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAG 418
DB 436 CTGTATCTTTAACTTCCCTCAGTCACTCTTTGGCAACGACCCCTGTCACATAAGA 495
QY 419 TGGCGGCGAGATCAAGAGGCGCTGTGGACACCGCGCGCGAGCAGCACCGTGTGGAGG 478
DB 496 TAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGGAGATGATACAGATTAGAAG 555
QY 479 AGATGACCTGCGCGGCAAGTGGGAAGCCCAAGATGATCGCGGATCGCGGCTTTCATCA 538
DB 556 AATAGATTGGCCAGGAAGATGGAACCAAAATGATAGGGGGAATTTGAGGTTTATCA 615
QY 539 AGGTGCGCCAGTAGCAGCATCTGTATCGAGATCTGCGGCAAGAGGCGCATCGGCACCG 598
DB 616 AAGTAAGACAGTATGATCAGATCTCATAGAACTGTGGACATAGAGCTATAGGTACAG 675
QY 599 TGTGATCGCGGCGACCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT 658
DB 676 TATTAGTAGGACCTTACACCTGTCAACATAATTTGGAAGAAATCTGTTCACTCAGATTGGTT 735
QY 659 GCACCTGMACTTCCCATCAGCCCATCGAGACCGTGGCCGTGAAAGCTGAAGCCCGGCA 718
DB 736 GCACCTTTAAATTTTCCCATTAGCCCTATTGAGACTGTACCAGTAAAAATTAAAGCCAGAA 795

QY 719 TGGACGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAAGATCAAGGCCCTGACCG 778
DB 796 TGGATGGCCCAAAAGTTAAACAATGGCCATTTGACAGAAGAAAAATAAAGCATTAGTAG 855
QY 779 CCATCTCGGAGGAGATGGAAGAGGAGGCAAGATCACCAAGATCGGCCCGGACCCCT 838
DB 856 AAATTTGTACAGAGATGGAAGAGGAGGCAAAATTTTCAAAATTTGGCCCTGAAATTCAT 915
QY 839 ACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTG 898
DB 916 ACAATACTCCAGTATTTGCCATAAAGAAAAAGACAGTAAATGGAAGAAATTAGTAG 975
QY 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGAGAGTGCAGCTGGGCATCCGCC 958
DB 976 ATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCAC 1035
QY 959 ACCCGCGCGCTGAAAGAAGAAGAGCGTGAACGCTGTGACGCTGGCGGACGCGCTACT 1018
DB 1036 ATCCCGCAGGCTTAAAAAAGAAAAATCAGTAACAGTACTGATGTGGTGATGATATT 1095
QY 1019 TCAGGTGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACATCCCGCATCA 1078
DB 1096 TTTCAAGTTCCCTTAGATGAAGACTTCAGGAAGTACTGCAATTTACCATACCTAGTATAA 1155
QY 1079 ACAACGAGACCCCGGCATCCGCTACCAAGTACAAAGTGTGCGCCAGGCTGGAAGGCA 1138
DB 1156 ACAATGAGACACAGGATTTAGATATCAGTACAATGTCTCCACAGGATGGAAGGAT 1215
QY 1139 GCCCGAGCATCTCCAGAGCAGCATGACCAAGATCTCTGAGCGCTTCCGCGCCGCAACC 1198
DB 1216 CACGAGCAATATTCAAAGTAGCATGACAAAAATCTTAGAGCCTTTATAGAAAAACAATC 1275
QY 1199 CCGAGATCGTATCTACCAGTACATGAGGACCTGTAGCTGGCGAGCACCTGGAGATCG 1258
DB 1276 CAGACATAGTTATCTATCAATACATGAGTATTTGATGTAGGATCTGACTTAGAATAG 1335
QY 1259 GCCAGCAGCGCCCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGCTTCAACA 1318
DB 1336 GGCAGCATAGAACAAAAATAGAGGAGCTGAGACACATCTGTTGAGTGGGAGCTTACCA 1395
QY 1319 CCCCAGACAGAAAGCAACAGAGGAGCGCCCTTCTGTGTGATGGGTACGAGCTGACCC 1378
DB 1396 CACGAGACAAAAACATCAGAAAGAACCTCCATCTCTTTGGATGGGTATGAACCTCCATC 1455
QY 1379 CCGACAGTGCACGCTGCAGCCCATCGAGCTGCCCCAGAAAGAGAGCTGGACCGTGAACG 1438
DB 1456 CTGATAATGGGAGTAGACAGCTATAGTGTGCGCAGAAAAAGACAGCTGGGACTGTCAATG 1515
QY 1439 ACATCCAGAAAGCTGGTGGGCAAGCTGAACTGGGCGCAGCAGATCTACCCCGCATCAAG 1498
DB 1516 ACATACAGAGTTAGTGGGAAATTTGAATTGGGCAAGTCAGATTTACCAGGGATTAAAG 1575
QY 1499 TGGCGCAGCTGTGCAAGCTGCTGCGCGCGCAAGGCGCTTACCGACATCGTGCCCTGA 1558
DB 1576 TAAGGCAATTTATGTAATCTTAGAGGAACCAAGCACTAACAGAACTAATACCACATA 1635
QY 1559 CCGAGAGGCGGAGCTGGAGCTGGCCGAGAACCGGAGATCTTGGCGAGCGCGCTGCACG 1618
DB 1636 CAGAAGAGCAGAGCTAGAACTGGCAGAAAAACAGAGATTTCTAAAAGAACCTGATACATG 1695
QY 1619 GCGTGTACTACGACCCCGCAGCAAGGACCTGTTGGCGGAGATCCAGAAGCAGGCGCACGACC 1678
DB 1696 GAGTGTATTATGACCCCATCAAAAGACTTAATAGCAGAAATACAGAGCAGGGGANGGCC 1755
QY 1679 AGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1738
DB 1756 AATGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAACAGAAAAATATGCAA 1815
QY 1739 AGATGCGCGCGCCCGCACCAACGCTGGAAGCAGCTGACCGAGCGCGCTGCAGAGATCG 1798
DB 1816 GAATGAGGGGTGCGCCACACTAATGATGTAACCAATTAACAGAGGAGCTGCAAAAAATAA 1875
QY 1799 CCATGGAGAGCATCGTGTCTGGGCAAGACCCCAAGTTCCGCTGCCCTCCAGGAAG 1858

```

1876 CCACAGAAAGCATAGTAATATGGGAAGAGACTCCTAAATTTAAACTGCCCATACAAAGG 1935
1859 AGACCTGGGAGACCTGGTGGACGACTACTGGCAGCGCCACCTGGATCCCGGAGTGGAGT 1918
1936 AAACATGGGAACATGGTGACAGAGATTGGCAAGCCACCTGGATTCTCTGAGTGGAGT 1995
1919 TCGTGAACACCCGCCCTGGTGAAGCTGTGTACCAGCTGGAGAAGGAGCCCATCATCG 1978
1996 TTGTTAATACCCCTCCTTAGTGAAATATGTTACAGTTAGAGAAGAACCCATAGTAG 2055
1979 GCGCCGAGACCTTCTACGTGGAGCGCGCCGCAACCGCGAGAGACCAAGATCGGCAAGCGG 2038
2056 GAGCAGAAACCTTCTATGTATGATGGGCGAGCTAACAGGAGAGACTAAATTAGGAAAAGCAG 2115
2039 GCTACCTGACCGACCGCGCGCAGCAAGATCGTGAGCCTGACCGGAGACCAACACAGA 2098
2116 GATATGTTACTAATAGAGGAACAAAGTGTGTACCCCTAATGACACAAACAAATCAGA 2175
2099 AGACCGAGTGCAGGCGCATCCAGCTGGCCCTGCAGGACACGCGGAGGTGAACATCG 2158
2176 AGACTGAGTTACAAGCAATTTAATAGCTTTGCGAGGATTCGGGATTAGAAGTAAACATAG 2235
2159 TGACCGACAGCAGTAGTACGCCCTGGGCATCATCCAGCGCCGACGCGGACAGAGCGAGCG 2218
2236 TAACAGACTCACAAATATGCAATTAGGAATCATTTCAAGCACAAACCCAGATCAAAAGTGAATCAG 2295
2219 AGCTGTGAACACAGATCAGCAGCAGCTGATCAAGAGGAGAGAGGTGTACCTGAGCTGGG 2278
2296 AGTTAGTCAATCAATTAATAGACAGTGTATAAAAAGGAAAAGGTCTATCTGGCATGGG 2355
2279 TGCCCGCCACAAAGGCGATCGCGCGCAAGCAGAGCAGATCGACAAAGCTGGTGAGCAAGGGCA 2338
2356 TACCAGCACAAAGGAATTTGGAGGAATGAACAAGTAGATAAATAGTCAGTGTCTGGAA 2415
2339 TCCGCAAGGTGCTTCTCGGAGCGGATCGAT 2370
2416 TCAGAAAGTACTATTTTAGATGGAATAGAT 2447
```

RESULT 3

```

US-09-117-217-11
; Sequence 11, Application US/09117217
; Patent No. 6221578
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117,217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (453)..(749)
; OTHER INFORMATION: Protease
US-09-117-217-11
```

```

Query Match 45.8%; Score 1132; DB 4; Length 2601;
Best Local Similarity 68.5%; Pred. No. 8.6e-178;
Matches 1626; Conservative 0; Mismatches 730; Indels 16; Gaps 4;

QY 14 TGGCCGAGCCATGAGCCAGGCCACCA---GCCCAACATCTCTGATGCGAGCGCAGCAACT 70
DB 77 TGGCTGAAGCAATGAGCCAAAGTAACAAATTCAGTACCATAATGATGCGAGAGGCAATT 136
```

```

QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGCAAGGAGGCCACATCGCCC 130
DB 137 TTAGGAACCAAAAGAAAGATTGTTAAGTGTTCATTTGTGCAAAAGAGGCGCACAGCCA 196
QY 131 GCAACTGCGCGCGCCCCCGCAAGAGGGTGTCTGGAAGTGGCGCAAGAGGGCCACACAGA 190
DB 197 GAAATTCAGGGGCCCCCTAGGAAAAAGGGCTGTGTTGAAAATGTGGAAGAGGAGACACAAA 256
QY 191 TGAAGACATGCACCGAGCGCCAGCCAACTTCTCCCGGAGGACCTGGCCCTTCCCCCAGG 250
DB 257 TGAAGATTGTACTGAGAGACAGGCTAA--TTTTTTAGGGAAGATCTGGCCCTTCTTACAAG 315
QY 251 GAAGGCCCGCGAGTTCCTCCCGAGCAGCAAAACCGCGCAACAGCCCAACAGCCACAGCGCGAGC 310
DB 316 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGACAGAGCCACAGCCCCACCAAGAAGAGC 375
QY 311 TGCAGTGGCGCGG-----CGAACACCCCGCAGCGAGGCGCGCGCGAGCGCCAGGGCA 364
DB 376 TTCAGGTCTGGGTAGAGACACAACACTCCCTCTCAGAAGCAGGAGCGCATAGACAAGGAA 435
QY 365 -----CCCTGAACCTTCCCCCAGATCACCTGTGGCAGCGCCCCCTGGTGAGCATCAAG 418
DB 436 CTGTATCTCTTAACTTCCCTCAGGTCACTCTTTGGCAACAGCCCCCTCGTCACAATAAGA 495
QY 419 TGGCGCGCCAGATCAAGAGGCGCCTGTGCACACCGCGCGCGAGCACACCGTGTGGAGG 478
DB 496 TAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 555
QY 479 AGATGAGCTGCGCGGCAAGTGGAAAGCCCAAGATGATCGCGCGCGATCGCGCGCTTCATCA 538
DB 556 AAATAGTGTTCGCAAGAGATGGAAACCAAAATGATAGGGGAATTTGAGGTTTTATCA 615
QY 539 AGTGCGCCAGTAGCAGCAGATCCTGATCGAGATCTCGCGCAAGAGGCCATCGGCACCG 598
DB 616 AAGTAAGACAGTATGATGAGATCTCATAGAAATCTGTGGACATAAAGCTATAGGTACAG 675
QY 599 TGCTGATCGGCGCCACCCCGTGAACATCATCGCGCGCGCAACATGCTGACCCAGCTGGGCT 658
DB 676 TATTAGTAGGACCTACACCTGTCAACATAATTGSAAGAAATCTGTCACCTCAGATTGGTT 735
QY 659 GCACCTGAACTTCCCGCATCAGCCCATCGAGACCGTGGCCCGTGAAGCTGAAGCCCGGCA 718
DB 736 GCACCTTAAATTTTCCCATTTAGCCCTATTGAGACTGTGTACCAAGTAAATTTAAGCCAGGAA 795
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
DB 796 TGGATGCGCCCAAGGTTAAACAATGGCCATTGACAGAAGAAAAAATAAAGCATTAGTAG 855
QY 779 CCATCTGCGAGGAGATGGGAAGGAGGCGTGAAGATCAACAAAGATCGGCCCGCGAGAACCCCT 838
DB 856 AAATTTGTACAGAGATGGAAGAAAGAGCGAAATTTTCAAAATTTGGCCCTGAAAATCCAT 915
QY 839 ACAACACCCCGTGTTCGCCCATCAAGAAGAGACAGACCAAGTGGCGGCAAGCTGGTGG 898
DB 916 ACAATATCTCCAGTATTTGGCCATAAGAAAAAAGACAGTACTATAATGGAGAAAAATTAGTAG 975
QY 899 ACTTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGCGAGCTGCGGATCCCC 958
DB 976 ATTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCAC 1035
QY 959 ACCCGCGCGCCCTGAAGAAGAAAGAGCGGTGACCCGTGCTGGACGTGGCGGCGAGCCCTACT 1018
DB 1036 ATCCCGCAGGCTTAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCAATATT 1095
QY 1019 TCAGCTGCGCCCTGGAGAGGACTTCGGCAAGTACACCGCTTCACCATCCCGAGCATCA 1078
DB 1096 TTTCAAGTTCCTTTAGATGAAGACTTCAGGAAGTATATCTGCAATTTTACCATACCTAGTATA 1155
QY 1079 ACAACAGAGACCCCGCATCCGCTACAGTACAACAGTGTGCTGCCCCAGGCGTGAAGGGCA 1138
DB 1156 ACAATGAGACACCCAGGATTTAGATATCAGTACATATGTGCTTCCACAGGGATGGAAGGAT 1215
```

QY 1139 GCCCAGCATCTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCGGCCCGCGCAACC 1198
Db 1216 CACCAGCAATATTCAAAAGTAGCATAACAAAATCTTAGAGCCCTTTAGAAAAACAATC 1275
QY 1199 CCGAGATCGTGATCTACCAAGTACATGACGACCTGTAGTGGGAGGAGCCTGGAGATCG 1258
Db 1276 CAGACATAGTTATCTCAATACATGATGATGTTGTAGTAGTCTGACTAGAAAATAG 1335
QY 1259 GCCAGCCGCGGCAAGATCCAGGAGCTGCGCAAGCAGCTGCTCGCTGGGGCTTCACCA 1318
Db 1336 GGCAGCATAGACAAAATAGAGAGCTGAGACACATCTCTTGGAGTGGGACTTACCA 1395
QY 1319 CCCCGCAAGAACACACAGAGAGGCCCTTCCTGTGGATGGGTACAGCTGCACC 1378
Db 1396 CACCAGACAAAACATCAGAAAGAACCTCCATTCTTTGGATGGTTATGAACCTCCATC 1455
QY 1379 CCGACAGTGGACCGTGGAGCCCATCGAGCTGCCGAGAGAGAGAGCTGGACGTGAACG 1438
Db 1456 CTGATAATGGACAGTACAGCCTATAGTCTGCCAGAAAAAGACAGCTGGACTGTCAATG 1515
QY 1439 ACATCCAGAGCTGTGGGCAAGCTGAACCTGGCCAGCCAGATCTACCCCGCATCAAG 1498
Db 1516 ACATACAGAAATAGTGGGAAATGAAATGGGCAAGTCAATTTACCCAGGGGATTAAG 1575
QY 1499 TGCGCCAGCTGTCAAGTGTGCGCGCGGCAAGGCCCTGACCGACATCTGTCGCCCTGA 1558
Db 1576 TAAGCAATTTATTAACCTCTTAGAGAACCAACAGCACTAACAGAGTAATACCACTAA 1635
QY 1559 CCGAGAGGCGGAGCTGGAGCTGCCGAGAACCGCGAGATCCTCGGGAGGCCCTGACG 1618
Db 1636 CAGAAAGAGCAGAGCTAGAAGTGGCAGAAAAACAGAGAGATTTCAAAGAACCACTACATG 1695
QY 1619 GCGTGTACTACGACCCAGCAGACGCTGGTGGCGGAGATCCAGACAGGCGCACGACC 1678
Db 1696 GAGTGTATTATGCCCATCAAAAGACTTAATAGCAGAAATACAGAGCAGGGGCAAGGCC 1755
QY 1679 AGTGAGCCTTACCAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGCGCAAGTACGCCA 1738
Db 1756 AATGGACATATCAAAATTTATCAAGGCAATTTAAATCTGAAAACAGAGAAATATGCAA 1815
QY 1739 AGATGCGGACCGCCACACCAACGCTGAAGAGCTGACCGGAGCGCGTGCAGAGATCG 1798
Db 1816 GAATGAGGGGTGCCACACTAATGATGTAAACAAATTAACAGAGCGAGTGCAAAATAA 1875
QY 1799 CCATGGAGAGCATGCTGATCTGGGCAAGACCCCAAGTCTCCGCTGCCATCCAGAGG 1858
Db 1876 CCACAGAAAGCATAGTATATATGGGAAAGACTCTTAATTTAACTGCCCATACAAAGG 1935
QY 1859 AGACTGGGAGACCTGTGTGACCGACTACTGGCAGGCGACCTGGATCCCCAGTGGAGT 1918
Db 1936 AAACATGSGAAACATGTGGACAGATATTGGCAAGCCACCTGGATTCTCGATGGGAGT 1995
QY 1919 TCGTGAACACCCCGCCCTGGTGAAGCTGTGGTACCGAGTGGAGAGGAGGCCATCATCG 1978
Db 1996 TTGTTAATACCCCTCCCTTAGTGAATTTATGGTACCAGTTAGAGAAAGAACCCCATAGT 2055
QY 1979 GCGCGGAGACCTTCTAGCTGAGCGCGCCGCAACCCGAGACCAAGATCGCAAGGCCG 2038
Db 2056 GAGCAGAAACCTTCTATGTAGTGGGAGCTAACAGGGAGACTAAATTTAGGAAGAGAG 2115
QY 2039 GCTACGTGACCGACCGGGCGGCGAGAGATCGTGAAGCTGACCGAGACCAACCAACAG 2098
Db 2116 GATATGTTACTAATAGAGGAGACAAAAGTTGTACCCCTAACTGACACACAATCAGA 2175
QY 2099 AGACCGAGCTCGAGGCATCCAGCTGGCCCTGCGAGGAGCGGCGAGGAGTGAACATCG 2158
Db 2176 AGACTGAGTTACAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAAGTAAACATAG 2235
QY 2159 TGACCGAGCGAGCTAGCCCTGGCATCATCCAGGCCCGCCGACAGAGCGAGAGCG 2218
Db 2236 TAAACAGACTCAAAATATGCAATAGGAATCAFTCAAGCACAAACAGATCAAAATCAG 2295
QY 2219 AGCTGGTGAACCAAGATCATCGAGCAGCTGATCAAGAGGAGAGGAGTGTACCTGAGCTGG 2278

Db 2296 AGTTAGTCAATCAATAATAGAGCAGTTAATAAAAAAGGAAAGTCTATCTGGCATGG 2355
QY 2279 TGCCCGCCCAAGGGCATCGCGGCAACGAGCAGATCGAAAGCTGTGAGCAAGGCA 2338
Db 2356 TACCAGCACACAGGAATTTGGAGAAATGAACAAATAGATAAATAGTCAGTGTGGAA 2415
QY 2339 TCCGCAAGGTGCTGTCTCTGGACGCATCGAT 2370
Db 2416 TCAGAAAGTACTATTTTATGATGAATAGAT 2447

RESULT 4
US-09-117-217-13
; Sequence 13, Application US/09117217
; Patent No. 6221578
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117,217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 13
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (750)..(2435)
; OTHER INFORMATION: Reverse Transcriptase
US-09-117-217-13

Query Match 45.8%; Score 1132; DB 4; Length 2601;
Best Local Similarity 68.5%; Pred. No. 8.6e-178;
Matches 1626; Conservative 0; Mismatches 730; Indels 16; Gaps 4;

QY 14 TGGCGGAGCCCATGAGCCAGGCGCACCA---GCGCCCAACATCTGATGCGCGCACCACT 70
Db 77 TGGCTGAGCAATGAGCCAAAGTACAAATTCAGCTACCATAATGATGAGAGGCAATT 136
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGGAGGCGCCACATCGCC 130
Db 137 TTAGAACCAAGAAAGATTGTTAAGTCTTTCAATTGTGGCAAGGAGGCGCACAGCCA 196
QY 131 GCAACTGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGCGGCAAGGAGGCGCCACCAGA 190
Db 197 GAAATTGCGAGGCGCCCTAGGAAAAAGGGCTGTTGGAATGTGGAAGGAAAGGACACCAA 256
QY 191 TGAAGGACTGCACCGAGCGCCGCAAGTCTTCTCCGCGAGGACCTGCGCTTCCCCCAGG 250
Db 257 TGAAGAATTGTACTGAGAGACAGGCTAA-TTTTATGGAAGATCTGGCTTCTCTACAG 315
QY 251 GCAAGGCGCGGAGTTCCTCCCGAGCAGAGAACCGCGCAACAGCCCGCCAGCGCGAGC 310
Db 316 GGAAGGCGAGGGAATTTTCTTCAGAGCAGACAGAGCCCAACAGCCCGCCAGCAAGAGAGC 375
QY 311 TCGAGGTGCGGG-----CGACAAACCCCGAGGAGCGCGCGCCGAGCGCCAGCGGCA 364
Db 376 TTCAGGTCTGGGGTAGAGACAACTCCCGCTCAGAGAGCAGGAGCGCGATAGACAGGAA 435
QY 365 -----CCTGAACCTTCCCGCAGATCACCTGTGCGACGCGCCCTGTTGAGCATCAAG 418
Db 436 CTGTATCTTTAACTTCCCTCAGGTCACTCTTTGGCAACAGCCCTCGTCAACAATAAGA 495
QY 419 TGGCGGCGCAGATCAAGGAGGCCCTGCTGGACACCGCGGCGGACACACCGCTGTGGAGG 478
Db 496 TAGGGGSCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATCATACAGTATTAGAG 555

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/418,848A
 FILING DATE: 07-APR-1995
 CLASSIFICATION: 526
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/936,849
 FILING DATE: 28-AUG-1992
 CLASSIFICATION: 526
 ATTORNEY/AGENT INFORMATION:
 NAME: RICHARD W. BORK
 REGISTRATION NUMBER: 36,459
 REFERENCE/DOCKET NUMBER: 2026-4091US2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-758-4800
 TELEFAX: 212-751-6849
 TELEX: 421792
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 7399 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-418-848A-9

Query Match 45.68; Score 1125.6; DB 2; Length 7399;
 Best Local Similarity 68.48; Pred. No. 1.1e-176;
 Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

```

QY 14 TGGCGAGGCGCATGAGCGAGCGCCACCA---CGCGCAACATCTGTATGCGAGCGCAGCAACT 70
Db 1838 TGGCTGAAGCAATGAGCGCAAGTACAAATCCAGCTACCATTAATGATACAGAAAGGCAATT 1897
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGTCTCAACTGGCGCAAGAGGGCCACATCGCCC 130
Db 1898 TTAGGAACCAAGAAAGACTGTTAAGTGTCTTCAATTTGGCANAAGAGGGGCACATAGCCA 1957
QY 131 GCAACTGCGCGCGCCCGCGCAAGAGGGCTGTGGAAGTGGCGGCAAGGAGGGCCACAGA 190
Db 1958 AAAATTGCGAGGGCCCCATGAGAAAAGGGCTGTGGAAATGTGGAAGGAGGAGCACACAAA 2017
QY 191 TGAAGGACTGCACCGAGCGCGCAGCGCAACTTCTTCCGCGAGGACTGTGGCCTTCCCGCAGG 250
Db 2018 TGAAGGATTTGATGAGAGACAGGCTAA-TTTTGTAGGAGATCTGGCCTTCCCAAG 2076
QY 251 GCAAGGCGCGGAGTTCCCGCAGCGAGCAGAACCGCGCCCAACAGCCCAAGCGCGGAGC 310
Db 2077 GGAAGGCCAGGGGAATTTCTTTCAGAGCAGACACAGAGCCCAACAGCCCAAGAGAGAGC 2136
QY 311 TGCAGGTTGCGGG---CGACAAACCCCGCAGCGAGGCGCGCGCGCGCGCGCGCGCGCA 364
Db 2137 TTCAAGTTTGGGGAAGAGACAACAACTCCCTCTCAGAAAGGAGGAGCGGATAGACAAGGAA 2196
QY 365 -----CCCTGAACCTTCCCGCAGATCACTCTGTGGCAGCGCCCGCTGGTGAAGCATCAAG 418
Db 2197 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACATAAGA 2256
QY 419 TGGCGCGCGCAGATCAAGAGGCGCTGTGGACACCGCGCGCGAGCACCGCTGTGGAGG 478
Db 2257 TAGGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 2316
QY 479 AGATGAGCCTTCCCGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGGCTTCATCA 538
Db 2317 AATGAATTTGCCAGGAAGATGAAACCAAAATGATAGGGGGAATTTGAGGTTTATCA 2376
QY 539 AGGTGCGCCAGTACGACCAAGTCTGTATCGAGATCTGTGGCAAGAGGCCATCGGCACCG 598
Db 2377 AAGTAGGACAGTATGATCAGATCTATAGAAATCTCGGACATTAAGCTATAGGTACAG 2436
QY 599 TGCTGATGGCGCCACCCCGTGAACATCATCGCGCAACATGCTGACCCAGCTGGGCT 558
Db 2437 TATTTAGTAGACCTACACTGTCAACATTAATTTGGAAGAATCTGTGACTCAGATTGGCT 2496
QY 659 GCACCTTGAACCTTCCCATTCAGCCCATCGAGACCGTGCCTGAGCTGAAGCTGAAGCCCGGCA 718
Db 2497 GCACCTTAAATTTTCCCATTTAGTCTATTGAGACTGTACCACTMAATTTAAGCCAGAA 2556
QY 719 TGGAGGCGCCCAAGGTGAAGAGTGGCCCTTACCGAGGAGAGATCAAGGCGCCCTGACCG 778
Db 2557 TGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAGAAAGAAAATAAAGCAATTAGTAG 2616
QY 779 CCATCTGCGAGGAGTGGAGAAGGAGGCAAGATCACCAAGATCGCCCGCGAGACCCCT 838
Db 2617 AATTTGTACAGAAATGGAAGAGGAGGAAATTTTCAAAATTTGGCGCTGAAATTCAT 2676
QY 839 ACAACACCCCGGTGTTCCCATTCAAAGAAAGAGCAGCACCAGTGGCGCAAGCTGGTGG 898
Db 2677 ACAATACTCCAGTATTTGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAAATTAGTAG 2736
QY 899 ACTTCGCGAGCTGAACAAGCCACCCAGGACTTCTGGAGAGTGGAGCTGGGCGATCCCC 958
Db 2737 ATTTTCAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCA 2796
QY 959 ACCCGCGCGCTGAAAGAAAGAGAGCGTGCACCGTGTGACAGTGGCGAGCGCTACT 1018
Db 2797 ATCTGCGAGGTTAAACAGAAAAAATCAGTAACAGTACTGGATGTGGCGATGATATT 2856
QY 1019 TCAGGCTGCCCTGGAGGAGGACTTCCGCAAGTACACCGCTTCCACCTACCCAGCATCA 1078
Db 2857 TTTCACTTCCCTTAGATAAAGACTTTCAGGAAGTATACTGCAATTTACCATACCTAGTATAA 2916
QY 1079 ACAGGAGACCCCGGCTACCGTACCAAGTACAAAGTGTGCTGCGCCAGGCTGGAGGCGCA 1138
Db 2917 ACAATGAGACACCGGATATAGATATAGTACAATGCTCTCCACAGGATGGAAAGAT 2976
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCGCGCGCCCAACC 1198
Db 2977 CACCAGCAATATTCAGTGTAGCATGACAAAAATCTTAGAGCCTTTTGAANAACAAATC 3036
QY 1199 CCGAGATCGTATCTACAGTACATGAGCAGCTGTAGCTGGCGAGCAGCTGGAGATCG 1258
Db 3037 CAGACATAGTATCTATCAATACATGAGTATTTGTATGTAGGATCTGACTTAGAATAG 3096
QY 1259 GCGAGCAGCGCGCAAGATCGAGGAGTGCAGCAGCAGCTGCTGCGCTGGGCGCTTACCA 1318
Db 3097 GGCAGCATAGAACAAAAATAGAGGAACCTGAGACACATCTGTTGAGGTGGGATTTTACCA 3156
QY 1319 CCGCCGACAAAGAACCCAGAGGAGCGCCCTTCCCTGTGATGGCTACGAGCTGCAC 1378
Db 3157 CACCAGACAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTTATGACTCCATC 3216
QY 1379 CCGACAGTGAACCGTGGAGCCCATTCAGCTGCGCGAGAGAGAGCTGGACCGTGAACG 1438
Db 3217 CTGATAATGACAGTACAGCTTATAGTCTGCCAGAAAGACAGCTGGACTGTCAATG 3276
QY 1439 ACATCCAGAGCTGTGGGCAAGCTGAACCTGGCGCAGCAGCATCTACCCCGGCATCAGG 1498
Db 3277 ACATACAGAAATAGTGGGAAAATTTGAATTTGGCAAGTCAAGATTTATGCGAGGATTAAG 3336
QY 1499 TGCCCGCAGCTGTGAAGCTGTGCGCGCGCGCAAGGCGCTGACCGACATCTGCGCCCTGA 1558
Db 3337 TAAGGCAATTATGTAACCTTTAGGGGAACCAAGACACTAACAGAGTACTACCACTAA 3396
QY 1559 CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCCTGCGGAGCGCCCTGACG 1618
Db 3397 CAGAAGAGCAGAGCTAGAACTGGCAGAAAAACAGGGAGATTTCTAAAAAACCCGGTACATG 3456
QY 1619 GCGTGTACTAGACCCCGCAGCAAGGACCTGTGCGCGGAGATCCTGCGGAGCGCCCGACGACC 1678
Db 3457 GAGTGTATTTAGCCCATCAAAAGACTTAATAGCAGAAATACAGAAAGCGGGGCAAGGCC 3516
QY 1679 AGTGGACCTACAGATCTACAGGAGCGCCTTCAAGAACCTGAAGACCGGCAAGTACGCA 1738

```

Db	3517	ATATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAACAGAGAAAATATGCAA	3576
QY	1739	AGATGGCGACCGCCACACCAACGACGCTGAAGCAGCTGACCGAGGCGGTGCAGAAGATCG	1798
Db	3577	GAATGAGGGTCCCACTACTATGATGTGAAACAATTAACAGAGGCAGTACAAAAATAG	3636
QY	1799	CCATGGAGAGCATCGTATCTGGGCAAGACCCCCAAAGTTCCGCTGCCCATCCAGAAGG	1858
Db	3637	CCACAGAAAGCATAGTAATATGGGGAAGAGACTCTAAATTTAAATTAACCATACAAAAGG	3696
QY	1859	AGACCTGGGAGACCTGGTGGACCCGACTACTGGCAGGCCACCTGGATCCC CGAGTGGGAGT	1918
Db	3697	NAACATGGGAAGCATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTCAGTGGGAGT	3756
QY	1919	TCGTGAACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGCAAGAGGCGCCATCATCG	1978
Db	3757	TTGTCAATACCCCTCCCTTAGTGAAGTTATGGTACCAGTTAGACAAGAACCCATATAG	3816
QY	1979	GGCGCGAGACCTTTCACGTGGAGCGCGCCGCAACCGCGAGACAAGATCGGCAAGCGCG	2038
Db	3817	GAGCAAAACTTTCTATGTAGATGGGCGAGCCAAATAGGGAACATAATTAGGAAAAAGCAG	3876
QY	2039	GCTACTGACCAACCGGGCGGCAAGAGATGTGTAGCCTGACCGAGACACCAACCCAGA	2098
Db	3877	GATATGTAATGACAGAGGAAGACAAAAAGTTGTCCCTTAACGCGACACACAAATCAGA	3936
QY	2099	AGACCGAGCTGAGAGCCATCCAGCTGGCCCTGCAGGACAGCGGAGCGAGTGAACATCG	2158
Db	3937	AGACTGAGTTACANGCAATTCATCATGTCTTTCGAGATTTCGGGATTAAGAGTAACATAG	3996
QY	2159	TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCCGCAAGACGAGAGCG	2218
Db	3997	TGACGACTCACAATATGCAATTTGGGAATCATTTCAAGCACACACCCAGATAAGAGTGAATCAG	4056
QY	2219	AGCTGGTGAACAGATCATCGAGAGCTGTATCAAGAAGCAAGAGTGTACCTGAGCTGGG	2278
Db	4057	AGTTAGTCAGTCAAAATAATAGAGCAGTTTAATANAAGAGAAAAAGTCTACCTGGCATGGG	4116
QY	2279	TGCCCCCCACAAAGGSCATCGCGGCAACGAGCAGATCGACAAGCTGGTGTAGCAAGGCA	2338
Db	4117	TACCAGCACAAAGGAATTGAGGAAATGAACAAGTATAGTTGGGTTGCTCAGTCTCGAA	4176
QY	2339	TCGCAAGGTGCTGTTTCCTGGAGCGGCATCGAT	2370
Db	4177	TCAGGAAAGTACTATTATTTTAGATGGAATAGAT	4208

RESULT 6

```

US-08-188-583-5
; Sequence 5, Application US/08188583
; Patent No. 5851813
;
; GENERAL INFORMATION:
;
; APPLICANT: Destroiers, Ronald C.
;
; TITLE OF INVENTION: PRIMATE LENTIVIRUS VACCINES
;
; NUMBER OF SEQUENCES: 57
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 MB
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,583
; FILING DATE:
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:

```

QY 719 TGGAGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGGATCAAGGCCCTGACCG 778
Db 2596 TGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAGAGAAAAATAAAGCATTAGTAG 2655
QY 779 CCATCTGCGAGGAGATGAGAGAGGAGGCAAGATCACCAGATCGGCCCGGAGAGCCCT 838
Db 2656 AAAATTGTTACAGAAATGGAAGAGAGGAGAAATTTCAAAATTTGGGCTGAAATCCAT 2715
QY 839 ACAACACCCCGTGTCCGCCATCAAGAGAGGAGCAGCAGCAAGTGGCGCAAGCTGGTG 898
Db 2716 ACAATCTCCAGTATTTGCCATTAAGAAAAAGACAGTACTAAATGGAGAAAAATTAGTAG 2775
QY 899 ACTTCGGGAGCTGAACAAGCCACCCAGGACTTCTGGAGGTGCGAGCTGGCGATCCGCC 958
Db 2776 ATTTACAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCAC 2835
QY 959 ACCCGCGCGCTGAAGAGAGAGAGAGCGTGAACGCTGCTGGAGTGGCGACGCTACT 1018
Db 2836 ATCTGCGAGGTTAAACAGAGAAAAATCAGTAACAGTACTGGATGTGGCGCATCATATT 2895
QY 1019 TCAGCGTCCCGTGGAGGAGACTTCCGCAAGTACACCGCCCTTACCATCCCGAGCATCA 1078
Db 2896 TTTTCAGTCCCTTAGATAAAGACTTCAGGAAGTATCTGCAATTTACCATCTAGTATAA 2955
QY 1079 ACACGAGACCCCGGATCCCGTACACAGTACACAGTGTGCTGCCCGAGGGCTGGAAGGCA 1138
Db 2956 ACAATGAGACACCGAGGATTTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAGGAT 3015
QY 1139 GCCCGAGCATCTCCAGAGCAGTACCAAGATCTTGGAGCCCTTCCGCGCCCGCAGC 1198
Db 3016 CACCAGCAATATCCAGTGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAAATC 3075
QY 1199 CCGAGATCGTATACAGTACAGTACAGTGTGCTGGCGAGGACCTGGAGATCG 1258
Db 3076 CAGACATAGTATCATCAATACATGATGATTTGTATGAGTATCTGACTTACAATAG 3135
QY 1259 GCGAGCAGCGGCAAGATCGAGAGCTGGCAGAGCTGCTGCGGTGGGGCTTACCA 1318
Db 3136 GGCAGCATAGAACAAAAATAGAGAACTGAGACAACTCTGTTGAGGTGGGGATTTACCA 3195
QY 1319 CCCCAGCAAGACGACAGAGGAGGCCCCCTTCTGTTGGATGGCTAGAGCTGAC 1378
Db 3196 CACCAGACAAAAACATCAGAGAAAGCTTCAATCTTTGGATGGTTATGAATCCATC 3255
QY 1379 CCGACAAAGTGGACGCTGACGCCATCAGCTGCCGAGAGAGAGCTGGACCGTGAACG 1438
Db 3256 CTGATAAATGACAGTACAGCTATAGTGTGCCAGAAAAAGCAGCTGGACTCTCAATG 3315
QY 1439 ACATCCAGAGCTGGTGGGCAAGCTGAAGTGGCGGAGGAGATCTACCCGGCATCAAG 1498
Db 3316 ACATACAGAAATTAGTGGGAAAAATTTGAATTTGGGCAAGTCAAGATTTATGAGGATTAAG 3375
QY 1499 TGGCCAGCTGTGCAAGCTGTGCGCGCGCGCCAGGCGCTGACGACATGTCGCCCTGA 1558
Db 3376 TAGGCAATATGTAACTTTTAGGGGAACCAAGACATTAACAGAGTATGACCACTAA 3435
QY 1559 CCGAGAGCGGAGCTGGAGCTGGCCGAGAACCCGAGATCTCTGCGAGCGCGTGCACG 1618
Db 3436 CAGAAGAGCAGAGCTAGAACTGGCAGAAAAACAGGGAGATTTAAAGAACCCGTACATG 3495
QY 1619 GCGTGTACTAGGACCCAGCAGAGGACCTGCTGGCGGAGATCCAGAACGAGGCGACGAC 1678
Db 3496 GAGTGTATTTAGCCCATCAAAAGACTTAATAGCAGAAATACAGAAAGCGGGCAAGGCC 3555
QY 1679 AGTGGACCTACAGATCTACGAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1738
Db 3556 AATGGACATATCAATTTATCAGAGCCATTTAAAAATCTGAACAGGAAAAATATGCAA 3615
QY 1739 AGATGGCACCCCAACCAACGAGTGAAGCAGCTGACCGGCGCTGCAGAAAGTGC 1798
Db 3616 GAATGAAGGTGCGCCACATTAATGATGTGAACAAATTAACAGAGGAGTCAAAAAATAG 3675
QY 1799 CCATGAGAGCATCGTATCTGGGCAAGACCCCAAGTTCCCGCTGCGCATCCAGAGG 1858

Db 3676 CCACAGAAAGCATAGTATATATGGGAAAGACTCTCTAAATTTAAATTTACCATACAAAAG 3735
QY 1859 AGACCTGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1918
Db 3736 AAACATGGGAAGCATGGTGGACAGAGTATTGGCAAGCCACCTGGATTTCTGAGTGGAGT 3795
QY 1919 TCGTGAACACCCCGCTGGTGGAGCTGTGTACCACTGGAGAGAGCCCATCATCG 1978
Db 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGGTACCAAGTATAGAGAAAGAACCCATAATAG 3855
QY 1979 GCGCGAGACCTTCTACGTGGACGCGCCCAACCGGAGACCAAGATCGGCAAGGCG 2038
Db 3856 GAGCAGAACTTTCTATCTAGATGGGCGAGCCAAATAGGAAACTAAATTAGGAAAGCAG 3915
QY 2039 GCTAGTGACCGACCGGGCGGCGAGAGATGTGTAGGCTGACCGAGACCAACCAAGCA 2098
Db 3916 GATATGTAACGAGAGAGAAAGTGTCCCTTAACGGACACAACAAATCAGA 3975
QY 2099 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTCCAGGACGCGGAGGAGTGAACATCG 2158
Db 3976 AGACTGATTTAAGCAATTCATCTAGCTTTGCGAGATTCGGGATTTAGAAAGTAAACATAG 4035
QY 2159 TGACCGACAGCAGTACGCCCTGGCGCATCTCCAGGCCAGCCGACAGAGCGAGAGCG 2218
Db 4036 TGACAGACTCAATATGCAATTTGGGAATCATTCAGACACACCAAGATGAATCAG 4095
QY 2219 AGCTGGTGAACAGATCATCGAGCAGCTGTATCAAGAGAGAGAGTGTACCTGAGCTGG 2278
Db 4096 AGTTAGTCAGTCAATAATATAGAGCAGTTAATAAAAAAGAAAAAGTCTACCTGGCATGG 4155
QY 2279 TGCCCGCCCAAGGCGCTGCGGCAAGCAGCAGCAGATCGCAAGCTGCTGAGCAAGGCA 2338
Db 4156 TACCAGCACACAAAGAAATTTGAGGAAATGAACAAATAGATGGTGGTGTGCTGTGAA 4215
QY 2339 TCCGCAAGGTGCTGTCTCTGGACGCGCATCGAT 2370
Db 4216 TCAGAAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 7
US-08-388-353-1
; Sequence 1, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: McPhee, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9709 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-388-353-1

Query Match 45.6%; Score 1125.6; DB 3; Length 9709;
Best Local Similarity 68.4%; Pred. No. 1.1e-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

QY 14 TGCCGAGGCGCATGAGCCAGGCGACCA---GCGCCAAACATCTCTGATGCGAGCGAGCAACT 70
DB 1877 TGCGTGAAGCAATGAGCCAGTAACAAATCCAGCTACCATAATGATACAGAAAGCAATT 1936

QY 71 TCAAGGGCCCCAAGCGGATCATCAAGTCTTCAACTGCGGCAAGGAGGCGGCACATCGCCC 130
DB 1937 TTAGGAACAAAGAAAGACTGTAAAGTGTTCAAATTGTGCAAGAAAGGCGACATAGCCA 1996

QY 131 GCAACTGCCCGCGCCCGGCAAGAGGCTGTGGAAGTGCAGCAAGGAGGCGCACAGA 190
DB 1997 AAAATTGACGGGCCCCTAGGAAAGAGGCTGTGGAATGTGGAAGGAAGGACACCAA 2056

QY 191 TGAAGACTGACCGAGCGCGCCAGGCAACTTCTTCCGCGAGGACCTTGGCCCTTCCCCCAGG 250
DB 2057 TGAAGACTGACCGAGCGCGCCAGGCAACTTCTTCCGCGAGGACCTTGGCCCTTCCCCCAGG 2115

QY 251 GCAAGGCCCGGAGTTCCTTTCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 310
DB 2116 GGAAGGCCCGGAGTTCCTTTCAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2175

QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGCGCGCGCGCGCGCA 364
DB 2176 TGCAGGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2235

QY 365 -----CCCTGAACTTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 418
DB 2236 CTGTATCTCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 2295

QY 419 TGGGCGCCAGATCAAGAGGCGCTTCTGTCAGACCGCGCGCGCGCGCGCGCGCGCGCG 478
DB 2296 TAGGGGGCAATTAAAGGAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAG 2355

QY 479 AGATGAGCTGCGCGGCAAGTGAAGCCCAAGATGATGCGCGGATCGCGCGGCTTCATCA 538
DB 2356 AAATGAATTTGCCAGGAAGATGAAACCAAAATGATAGGGGAAATTGGAGGTTTTATCA 2415

QY 539 AGGTGCGCAGTAGGACCATCTGATTCAGATCTGCGGCAAGAGGCGCATCGGCACCG 598
DB 2416 AAGTAGGACAGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2475

QY 599 TGCTGATGCGCGCCCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 658
DB 2476 TATTAGTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2535

QY 659 GCACCTTGAATTTCCCATCAGCCCATCAGAGCCGTCGAGGAGGAGGAGGAGGAGGAGG 718
DB 2536 GCACTTTAAATTTTCCCATTTAGTCTATTGAGAGTGTACAGTAAATTTAAAGCCAGGAA 2595

QY 719 TGNAGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 778
DB 2596 TGGATGGCCCAAGAGTAAACAAATGGCCATTGACAGAGGAGGAGGAGGAGGAGGAGG 2655

QY 779 CCATCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 838
DB 2656 AAATTGTACAGAAATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2715

QY 839 ACAACACCCCGGTTTCGCCATCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 898

DB 2716 ACAATACTCCAGTATTTCCTCAATAAGAAAAAGACAGTACTAATAAGGAGAAAAATTAGTAG 2775

QY 899 ACTTCGCGAGCTGAACAAGCGCAGCAGGACTTCTGGGAGGTGCGAGCTGGGGATCCCC 958

DB 2776 ATTTTCAGAGAACTTAATAAGAGAACTCAAGATTCTGGGAAGTTCAATTTAGSAAATCCAC 2835

QY 959 ACCCCGCGCGCTGAAGAAGAAGAGCGTGACCGTGTGAGAGCTGGGGGAGCGCTACT 1018

DB 2836 ATCTTCAGGGTTTAAACAGAAAAATCAGTAACAGTACTGGATGTGGGCGATGCATATT 2895

QY 1019 TCAGCTGCGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 1078

DB 2896 TTTTCAGTTCCTTTAGATAAGACTTCAGGAAGTATACATGCTTTACCATACCTAGTATAA 2955

QY 1079 ACAACGAGACCCCGGATCCGCTACAGTACAAGCTGTGCTGCCCCAGGCTGGAAGGGCA 1138

DB 2956 ACAATGAGACACCGAGGATTTAGATATCAGTACAATGTGCTTCCACAGGATGGAAGGAT 3015

QY 1139 GCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCGCAACC 1198

DB 3016 CACAGCAATATTCAGTGTAGATGACAAAAATCTTAGAGCCTTTTAGAAAAACAATC 3075

QY 1199 CCGAGATCTGTATCTACAGTACATGAGACACCTGTACGTGGGCGAGCGCTGGAGATCG 1258

DB 3076 CAGCATAGTCTATCTATCAATACATGATGATTTGTATGTAGTCTGACTTAGAAATAG 3135

QY 1259 GCAGACACCGCGCAGATCGAGAGCTGCGAAGCACTGTGCTGCGCTGGGGCTTCACCA 1318

DB 3136 GGCAGCATAGAAACAAAAATAGAGGAACTGAGACAACATCTGTTGAGGTGGGATTTACCA 3195

QY 1319 CCCCCACAAGAGCAGCAGAAAGGAGCGCCCTTCTCTGTGATGGGTAGGAGTGCAC 1378

DB 3196 CACGACAAAAACATCAGAAAGACCTCCATCTCTTGGATGGGTTATGAATCTCATC 3255

QY 1379 CCGACAAGTGGACCGTGCAGCCCATCGAGTGCCTGCGGAGAGGAGAGTGGACCGTGAACG 1438

DB 3256 CTGATAAATGGACAGTACAGCTTATAGTGTGCGCAAGAAAGACAGCTGGACTGTCAATG 3315

QY 1439 ACATCAGAGAGCTGGTGGCAAGCTGAACTGGGCGCAGCTACCTACCCGGCATCAAGG 1498

DB 3316 ACATACAGAAATTAGTGGGAAATTTGAATTTGGGCAAGTCAAGTTTATGAGGGATTAAG 3375

QY 1499 TGGGCCAGCTGTGCAAGCTGTGCGCGCGCCCAAGGCGCTGACCGACATCGTCCCGCTGA 1558

DB 3376 TAAGGCAATTTATGTAATCTTTAGGGGAAACCAAGCACTAACAGAAAGTAGTACCCTAA 3435

QY 1559 CCGAGAGGCGGAGCTGGAGCTGGCCCGAGAACCGCGAGATCTCTGCGGAGCGCCGTGCACG 1618

DB 3436 CAGAAGAGCAGAGCTAGAACTGGCAGAAACAGGGAGATTTCTAAAGAACCGGTACATG 3495

QY 1619 GCGTGTACTAGGACCCCGCAGCAAGGAGCTGTGGCGGAGATCCAGAAAGCAGGCGCCAGC 1678

DB 3496 GAGTGTATTATGACCCCATCAAGAGCTTATAGCAGAAATACAGAGCAGGGGCAAGGCC 3555

QY 1679 AGTGGACCTACAGATCTACAGGAGCGCTTCAAGAACCTTGAAGACCGCGCAAGTACGCCA 1738

DB 3556 AATGGACATATCAAAATTTATCAAGAGCCTTTAAATAATCTGAAAAACAGGAAAAATGCAA 3615

QY 1739 AGATGCCACCGGCCACACACGACGCTGAAGAGCTGACCGGAGCGCGTGCAGAGATCG 1798

DB 3616 GAATGAAGGGTGGCCACACTAATGATGTGAACAAATTAACAGAGGGCAGTACAAAAATAG 3675

QY 1799 CCATGGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCCCGCTTCCCATCCCAAGG 1858

DB 3676 CCACAGAAAGCATAGTAAATTTAGGAAAGACTCCTAAATTTAAATTTACCCATACAAAGG 3735

QY 1859 AGACTGGGAGACTGTGTGACCGACTACTGTGAGGCGACCTGTGATCCCCAGGTGGGAGT 1918

DB 3736 AAACATGGGAAGCATGTGTGACAGAGTATTGGCAAGCACCTGGATTCCTTGATGGGAGT 3795

QY 1919 TCGTGAACCCCCCTTGGTGAAGCTGTGTACGAGCTGGAGAGGAGGAGGAGGAGGAGG 1978

Db 3796 TTGTCATAACCCCTCCCTTAGTGAAGTTATGTTACCATGTTAGAGAAAGAACCCATAATAG 3855
QY 1979 GCGCCAGACCTTCTACGTGGAGCGCGCCCAACCGAGACCAAGATCGGAAGCGG 2038
Db 3856 GAGCAAAACTTCTATGTAGATGGGCGCCCAATAGGGAATAAATAGGAAAGCAG 3915
QY 2039 GTTACGTGACCGCGCGCGGAGAGATGCTGAGCTGACCGAGACCAACACAGA 2098
Db 3916 GATATGTAAGTACAGAGGAGCAAAAGTTGTCCCTTAAGGACACACAAATCAGA 3975
QY 2099 AGACCGAGTGCAGGCGCATTCAGCTGGCCCTGCAGGACAGCGCAGGTGAACATCG 2158
Db 3976 AGACTGAGTCAAGCAATTCATCTAGCTTTCAGGATTCGGGATTAGAAGTAACATAG 4035
QY 2159 TGACCCAGCAGTACGCGCTGGGCATCATCCAGCCCGCCGAGGAGGAGCG 2218
Db 4036 TGACAGCTACAAATATGATTTGGGAATCATTCACGACAAACAGATAGAGTGAATCAG 4095
QY 2219 AGCTGTGAACCATGATCGAGCAGCTGATCAAGAGGAGAGAGTGTACCTGAGCTGGG 2278
Db 4096 AGTTAGTCAGTCAAAATATAGACAGTGAATATAAAGGAGGAAAGTCTACCTGGCATGGG 4155
QY 2279 TGCCCGCCCAAGGCGATCGGCGGCAAGCAGCAGATCGACAGCTGGTGAGCAAGGCGA 2338
Db 4156 TACCAGCACACAAAGGAATTTGGAGGAATGAACAAGTAGATGGGTTGGTCAGTGTGGAA 4215
QY 2339 TCCGCAAGTGTGTCCTGCGGCGCATCGAT 2370
Db 4216 TCAGGAAGTACTATTTTATAGATGGAATAGAT 4247

RESULT 8

US-08-488-551B-1
; Sequence 1, Application US/08488551B
; Patent No. 6015661
; GENERAL INFORMATION:
; APPLICANT: Nicholas J. Deacon
; APPLICANT: Dale A. McPhee
; APPLICANT: David Cooper
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 841
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,551B
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PM3864 (AU)
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: PM4002 (AU)
; FILING DATE: 21-FEB-1994
; APPLICATION NUMBER: PM0284 (AU)
; FILING DATE: 23-DEC-1994
; APPLICATION NUMBER: US 08/388,353
; FILING DATE: 14-FEB-1995
; APPLICATION NUMBER: PM3021/95
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FRANK S. DIGIGLIO
; REFERENCE/DOCKET NUMBER: 9606Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9709 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-488-551B-1

Query Match 45.6%; Score 1125.6; DB 3; Length 9709;
Best Local Similarity 68.4%; Pred. No. 1.1e-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

QY 14 TGCCCGAGGCGATGAGCGGCGCACCA---GCGCCAAATCCTGATCAGCGGCGCAACT 70
Db 1877 TGCTGAAGCAATGAGCCAAGTACAAATCCAGTACCATAATGATACAGAAAGCAATT 1936
QY 71 TCAAGGGCCCCCAGCGGATCATCAAGTGTTCAACTGCGGCAAGGAGGGCCACATCGCCC 130
Db 1937 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATATTGTGCAAAAGAGGCGCATAGCCA 1996
QY 131 GCAACTGCCGCGCCCCCGCAAGNAGGCTGCTGGAAGTGGCGGCAAGGAGGCGCACAGA 190
Db 1997 AAAATTGCGAGGGCCCCCTAGGAAAAGGGCTGTGGAAATGTGGAAGAGGAGACACCAA 2056
QY 191 TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGG 250
Db 2057 TGAAGATTGCTACTGAGAGACAGGCTAA-TTTTATGGGAAGATCTGCGCCTTCCCACAAG 2115
QY 251 GCAAGGCCCGCGAGTTGCCCGAGCGAGCAGAACCGCGCCACAGCCCCCAGCGCGGAGC 310
Db 2116 GGAAGGCCAGGGAATTTCTTTCAGAGCAGACAGAGCAACAGCCCCCAGGAGAGAGC 2175
QY 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGGCGCGCGGCGGCGAGGCA 364
Db 2176 TTAGGTTGGGGAAGAGACAACAACTCCCTCTCAGAAAGAGGAGCGGATAGACAAGAA 2235
QY 365 -----CCCTGAACCTTCCCCCAGATCACCTGTGCGAGCGCCCCCTGGTGAGCATCAAG 418
Db 2236 CTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTGTCACATAAAGA 2295
QY 419 TGGGCGGCCAGATCAAGAGGCGCCTGCTGGACACCGCGCGCGAGACCGGTGTGAGG 478
Db 2296 TAGGGGGCAATTAAAGGAAGCTCTATTAGATACAGAGCAGAGATGATACAGTATTAGAAG 2355
QY 479 AGATGACCTGCCCGCAAGTGGNAGCCCAAGATGATGCGGCGCATCGCGGCTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGGAAACCAAAATGATAGGGGGAATTTGGAGGTTTATCA 2415
QY 539 AGGTGCCCGCAGTACGACAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
Db 2416 AGTAGGACAGATGATGATCAGATCACTATAGAAATCTCGGACATAAAGCTATAGTACAG 2475
QY 599 TGCTGATCGGCGCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGTGGGCT 658
Db 2476 TATTAGTAGGACCTTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGAACTTCCCCATCAGCCCCATCGAGACCGTGGCCGTGAAGTGAAGCCCGCA 718
Db 2536 GCACCTTAAATTTTCCCATTTAGTCTCTATTGAGACTGTACCAAGTAAATTAAGCCAGGA 2595
QY 719 TGGACGGCCCCAAGGTGAAGCAGTGGCCCTGTACCGGAGGAGAAAGATCAAGGCCCTGACCG 778
Db 2596 TGGATGCCCAAAAGTTAAACAATGGCCATTTGACAGAAAGAAATTAAGCATTAGTAG 2655
QY 779 CCATCTCGAGGAGATGGAGAGGAGGCAAGATCAACAAGATCGGCCCGGAGAACCCCT 838
Db 2656 AAATTTGTACAGAAATGGAAGGAGGAAATTTTCAAAATTTGGGCTGAAATCCAT 2715
QY 839 ACACACCCCGGTTTCCGCATCAAGAGAGGACAGCACCAAGTGGCGCAAGCTGGTG 898
Db 2716 ACAATCTCCAGTATTTGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAAATAGTAG 2775

```
QY 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGCTGGGCATCCCC 958
Db 2776 ATTTCAGAGAACTTAATAAGAGAACTCAAGATTCTCGGGAAGTTCAATTAGGAATACCAC 2835
QY 959 ACCCCGCGGCGCTGAAGAAGAAAGAGAGCTGACCGTGTGGAGGTGGGGGACGCGCTACT 1018
Db 2836 ATCCTGCGAGGGTTAAAACAGAAAATACAGTACAGTACTGGATGTGGGGGATGCAATATT 2895
QY 1019 TCAGCGTGCCTCGGAGGAGCTTCCGCAAGTACACCGCTTCAACCATCCCGAGCATCA 1078
Db 2896 TTTCAAGTTCCTTAGATAAAGACTTCAGGAAGTATACGTGATTTACCTACCTAGTATAA 2955
QY 1079 ACAACGAGACCCCGGATCCGCTACAGTACAACTGTCTGCCAGGGGTGGAAGGGCA 1138
Db 2956 ACAATGAGACACCGAGGATTAGATATCAGTACAAATGTGCTTCACACAGGGATGAAAGGAT 3015
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGGNAAC 1198
Db 3016 CACCAGCAATATTCCAGTGTAGCATGACAAAATCTTAGAGCCCTTTAGAAAACAAATC 3075
QY 1199 CCGAGATCGTGATCTACCACTACATGGACGACCTGTACGTGGCAGGACCTGGAGATCG 1258
Db 3076 CAGCATAGTCACTCATCATACATGATGATTTGTATGTAGATCTGACTTGAANAATAG 3135
QY 1259 GCAGACCCCGGCGAAGATCGAGAGCTGCGCAAGCACTGTCTGCGGTGGGGTTCAACCA 1318
Db 3136 GCGAGCATAGAAAATAAGAGAACTGAGACAACATCTGTTGAGGTGGGATTTACCA 3195
QY 1319 CCGCCGACAGAAAGCACCAGAGGAGCCCTTCTCTGTGGATGGGCTAGCAGCTGGACC 1378
Db 3196 CACGACAAAATAACATCAGAAAAGACCTCCATCTCTTGGATGGGTATGAACCTCCATC 3255
QY 1379 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG 1438
Db 3256 CTGATTAATGGACAGTACAGCTTATAGTCTGCCAGAAAGACAGCTGGACTGTCAATG 3315
QY 1439 ACATCAGAAAGCTGTGGGCAAGCTGAAGTGGGCGCAGCCAGATCTTACCCCGGATCAAG 1498
Db 3316 ACATCAGAAATTAAGTGGGAAAATTTGAATTTGGCAAGTCAAGATTTATGAGGGATTAAG 3375
QY 1499 TGGCCAGCTGTCAAGCTGTCTCGCGGCGCCAAAGCCCTGACCGACATCGTCCCTGA 1558
Db 3376 TAAGGCAATTAATCTTTAGGGGAACCAAGACCTAACAGAAAGTAGTACCACATAA 3435
QY 1559 CCGAGGAGCGCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGCGAGCCCGTCAAG 1618
Db 3436 CAGAAGAAGCAGAGCTAGAAGTGGCAGAAAACAGGGAGATTTCTAAAGAAACCGGTACATG 3495
QY 1619 GCGTGTACTACGACCCAGCAGGACCTGGTGGCCGAGATCCAGAAAGCAGGCGCAGACC 1678
Db 3496 GAGTGTATTATGACCATCAAAAGACTTAATAGCAGAAATACAGAAAGCAGGCGCAAGGCC 3555
QY 1679 AGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGGAAGCCGCAAGTACGCCA 1738
Db 3556 AATGGACATATCAAAATTTATCAGAGCCATTTAAAATCTGAAAACAGGAAAATATGCAA 3615
QY 1739 AGATGCGCACCCGCCACACCAAGACGTGAAGCAGCTGACCGAGCGCGTGCAGAAATCG 1798
Db 3616 GAATGAAGGCTGCCACACTAATGATGTGAAACAATTAACAGAGGAGTACAAAATAAG 3675
QY 1799 CCATGGAGAGCATCTGTATCTGGGGCAAGACCCCAAGTTCGCCCTGCGCCATCCAGAAAG 1858
Db 3676 CCACAGAAAGCATAGTAATATGGGGAAGAGCTCTTAATTTAAATTTACCATCAAAAAG 3735
QY 1859 AGACCTGGGAGACCTGCTGGACGACTACTTGGCAGGCCACCTGGATCCCGAGTGGAGT 1918
Db 3736 AAACATGGGAGCATGCTGGACAGATATTGGCAAGCCACTGGATCTCTGAGTGGGAGT 3795
QY 1919 TCGTGAACACCCCGCCCTTGGTGAAGCTGTGATACAGCTGGAGAGGAGGCCCATCATCG 1978
Db 3796 TTGTCAATACCCCTCTCCTTAGTGAAGTTATGTTACCAATTTAGTGAAGAAAGAACCCATAATAG 3855
QY 1979 GCGCCGAGACCTTCTACGTGGAGGCGCGCCCAACCGCCGAGACCAAGATCGGCAAGCGCG 2038
```

```
Db 3856 GAGCAGAACTTCTCTATGTAGTGGGCGAGCCAATAGGAAACTTAATTAGGAAAGCAG 3915
QY 2039 GCTACGTGACCGACCGGGCGGCGAGAGATCGTGGAGCCTGACCGAGACCAACCAACAGA 2098
Db 3916 GATATGTAACTGACAGAGAGAACAAAAAGTTGTCGCCCTTAACGAGACAAACAATCAGA 3975
QY 2099 AGACCGAGCTGCAGGCCATCCAGCTGCCCTGACAGGACAGCGGCGAGGTGAACATCG 2158
Db 3976 AGACTGAGTTTACAAGCAATTCATCTAGCTTTGCAAGGATTCGGGATTAAGAAGTAACATAG 4035
QY 2159 TGACCGACACCGAGTACGCCCTTGGGCATCATCCAGGCGCCAGCCGACAAAGAGCGAGCG 2218
Db 4036 TGACAGACTCACAATATGTCATTTGGGAATTCATCAAGCACAACCAAGATGAATCAG 4095
QY 2219 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACTGTAGCTGGG 2278
Db 4096 AGTTAGTCACTCAAAATAATAGAGCAGTTAATAAAAAAGAAAAGTCTAGCTGGCATGG 4155
QY 2279 TGCCCGCCCAAGGGCATCGCGGCAACGAGCAGATCGACAAGCTGCTGAGCAAGGCA 2338
Db 4156 TACAGCACACAAAGGAATTTGGAGGAATGAACAAGTAGATGGTTGGTCAGTCTCGAA 4215
QY 2339 TCCGCAAGGTGCTGTTCTCTGACGCGCATCGAT 2370
Db 4216 TCAGGAAGTACTATTTTAGATGGAATAGAT 4247
RESULT 9
US-09-309-572-15
; Sequence 15, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; EARLIER FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 15
; LENGTH: 9709
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-309-572-15
Query Match 45.6%; Score 1125.6; DB 4; Length 9709;
Best Local Similarity 68.4%; Pred. No. 1.1e-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;
QY 14 TGCCCGAGCCCATGAGCCAGGCCACCA---GCGCCACATCTCTGTGTCAGCGCAGCACT 70
Db 1877 TGCTGAAGCAATGAGCCCAAGTAACAAATCCAGCTACCAATAATGATACAGAAAGCAATT 1936
QY 71 TCAAGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGCAAGGAGGCGGCACATCGCCC 130
Db 1937 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATTTGTGCAAGAGAGGCGCATAGCCA 1996
QY 131 GCAACTGCCCGCCCCCGCAAGAGGGCTGTGGAAGTGCAGGAGGCGGCGCCACAGA 190
Db 1997 AAAATTGAGGGCCCCCTAGGAAAAGGGCTGTTGGAATTTGTGAAAAGGAGGACACCAA 2056
QY 191 TGAAGACTGACCCGAGCGCCAGGCCCAACTTCTTCGCGAGGACCTTGGCTTCCCCCAGG 250
Db 2057 TGAAGATTGTACTGAGAGACAGGCTTAA-TTTTTTAGGGAAGATCTGGCTTCCCCAAG 2115
QY 251 GCAAGGCCCGCGAGTTCCCGCAGGACGACCGCCCAACAGCCACCCACCCAGCGGAGC 310
Db 2116 GGNAGGCCAGGGAATTTTCTTCAAGACACAGACAGCCACCCACCCAGAGAGAGC 2175
QY 311 TCGAGTGGCGG-----CGACAACCCCGCGCAGCGAGCGCGCGCGCGAGCGCGGCA 364
```

Db 2176 TTTGAGGTTTGGGGAAGAGACACAACTCCCTCTCTGAGAAGCAGGAGCCGATAGACAAGGAA 2235
QY 365 -----CCCTGAACCTTCCCCAGATCACCTCTGTGCGAGCGCCCTGGTGTGACATCAAGG 418
Db 2236 CTGTATCCCTTACCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAAAATAAGAA 2295
QY 419 TGGGGGGCCAGATCAAGGAGGCCCTGCTGGACACCGGGCCGACGACACCGCTGCTGGAGG 478
Db 2296 TAGGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAG 2355
QY 479 AGATGAGCTGCCCGGCAAGTGAAGCCCAAGATGATGGCGGCGATCGCGGCGCTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAGATGGAACCAAAATGATAGGGGGAATTTGGAGGTTTTATCA 2415
QY 539 AGGTGGCGGATPACGACCAATCTGATCGAGATCTGGGGCAAGAGGCCATCGGCACCG 598
Db 2416 AAGTAGGACAGTATGATCAGATCTCATAGAATCTGGGACATAAGAGCTATAGGTACAG 2475
QY 599 TGCTGATCGGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCT 658
Db 2476 TATTAGTAGGACCTACACCTGTCAACATAAATTGGAAGAAATCTGTGACTCAGATTGGCT 2535
QY 659 GCACCTTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCA 718
Db 2536 GCACCTTGAATTTCCCATTTAGTCTTATTTGAGACTGTACCAAGTAAATTAAGCCGAGGAA 2595
QY 719 TGGAGGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGAACCG 778
Db 2596 TGGATGGCCCAAGGTTAAACATGGCCATTTGACAGAGAAATAAAGCATTTAGTAG 2655
QY 779 CCATCTGCGAGAGATGGAGAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT 838
Db 2656 AAATTTGTACAGAAATGGAAGGAGGAAATTTCAAAATTTGGCCTGAAATTCAT 2715
QY 839 ACAACACCCCGTGTTCGCCATCAGAGAGAGGAGCAGCACCAGTGGCGCAAGCTGGTG 898
Db 2716 ACAATCTCCAGTATTTGCCATTAAGAAAGAAACAGTACTAAATTTGGAGAAATTTAGTAG 2775
QY 899 ACTTCGCGAGCTGAACAGGCGACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC 958
Db 2776 ATTTTCAGAGAACTTAATAAGAGACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCAC 2835
QY 959 ACCCGCGCGGCTGAGAGAGAGAGACGCTGACCGTCTGGAGCTGGGCGACGCTTACT 1018
Db 2836 ATCTTGCAGGTTTAAACAGAGAAATCAGTAACAGTACTGATGTGGCGGATGCTATTT 2895
QY 1019 TCAGGCTGCCCTGGACGAGGACTTCCCAAGTACACCGCTTCACCATCCCGACATCA 1078
Db 2896 TTTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATCTGCAATTTACCATCTAGTATAA 2955
QY 1079 ACAAGGAGACCCCGGATCCGCTACCAAGTACAACGTGCTGCCCGAGGGCTGGAAGGCA 1138
Db 2956 ACAATGAGACACCGGATTTAGATATCAGTACATGCTCTCCACAGGGATGGAAGGAT 3015
QY 1139 GCCCGAGCATCTCCAGAGCAGATGACCAAGATCTCGGACCCCTTCGCGCCCGCAACC 1198
Db 3016 CACGAGCAATTTCCAGTGTAGCATGACAAATAATCTTAGAGCCCTTTAGAAAACAAATC 3075
QY 1199 CCGAGATCTGATCTACAGTACATGAGCAGCTGTAGTGGCGAGCGACCTGGAGATCG 1258
Db 3076 CAGACATAGTCTATCAATACATGATGATTTGTATGTAGGATCTGACTTAGAAATAG 3135
QY 1259 GCCAGACCGGCGCAAGATCGAGGAGCTGGCGAAGCAGCTCTGCTGGTGGGCTTCACCA 1318
Db 3136 GGCAGCATAGAAAATAAGAGCACTGAGCAACATCTCTGTAGGTGGGATTTACCA 3195
QY 1319 CCCCAGACGAGCAACGACGAGGAGCCCTTCCTCTGTGGATGGGTACAGCTGCACC 1378
Db 3196 CACGAGCAAAAAACATCAGAAAAGAACTCCATTCCTTTGGATGGGTATGAATCCATC 3255
QY 1379 CCGACAAGTGACCGTGCAGCCCATCGAGCTGCCCGAGAGAGGAGCTGGACCCGTGAACG 1438

Db 3256 CTGATAAATGGACAGTACAGCCTATAGTCTGCCAGAAAGACAGCTGGAGTCTCAATG 3315
QY 1439 ACATCCAGAAGCTGTGGGCAAGCTGAATGGGCGCAGCAGATCTACCCCGGATCAAGG 1498
Db 3316 ACATACAGAAATTAGTGGGAAATTAATTTGGSCAAGTCAGATTTATGCAAGGATTTAAAG 3375
QY 1499 TCGGCCAGCTGTGCAAGCTGCTCGCGGGCCCAAGGCCCTGACCGGACATCTGTCGCCCTGA 1558
Db 3376 TAAGCAATTTATGTAACCTTTAGGGGAACCAAGACACTAACAGAAGTAGTATCCACTAA 3435
QY 1559 CCGAGGAGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGCGAGCCCGTGCACG 1618
Db 3436 CAGAAGAAGCAGAGCTAGAACTGGCAGAAACAGGAGATCTCTAAAGAAGACCGGTACATG 3495
QY 1619 GCGTGTACTACGACCCCAAGGACCTGGTGGCCGAGATCAGAGAAGCGGCGCCAGACC 1678
Db 3496 GAGTGTATTTATGACCCCTCAAAAGACTTAATAGCAAAATACAGAAGCAGGCGCAAGGCC 3555
QY 1679 AGTGACCTTACCAGATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCCGCAAGTACGCCA 1738
Db 3556 AATGGACATATCAATTTATCAGAGCCATTTAAANAATCTGAAACAGGAAATATGCAA 3615
QY 1739 AGATGGCAGCCGCCACACACGACGCTGAAGACGCTGACCGAGCGCGTGCAGAAAGTCG 1798
Db 3616 GAATGAAGGTTGCCACACTAATGATGTGAACAAATTAACAGAGCGAGTACAAAAATAG 3675
QY 1799 CCATGGAGAGCATCTGTGATCTGGGCAAGACCCCAAGTTTCGCCCTGCCCTCCAGAAAG 1858
Db 3676 CCACAGAAGCATAGTAAATATGGGGAAGACTCTTAATTTAAATTTACCCATACAAAAG 3735
QY 1859 AGACCTGGGAGACCTTGTGGACCGCTACTGTGCAGGCCACCTGGATCCCGAGTGGGAGT 1918
Db 3736 AACAATGGAGCATGTTGGACAGAGTATTGGCAAGCCACCTGGATTTCTGTAGTGGAGT 3795
QY 1919 TCGTGAACACCCCGCTGTGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCG 1978
Db 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGTTACAGTTAGAGAAAGAACCCATAATAG 3855
QY 1979 GCGCGAGACCTTCTACGTGGAGCGCGGCCCAACCGGAGACCAAGATCGGCAAGGCCG 2038
Db 3856 GAGCAGAATCTTCTATGTAGATGGGCGAGCCATAGGGAACCTAAATTTAGGAAAAGCAG 3915
QY 2039 GCTAGTCAACGACCGGGCGCGCAGAGATCGTGAAGCTGACCGCTGACCGACCAACCAAG 2098
Db 3916 GATATGTAACGTACAGAGAGAGACAAAAAGTTGTCCCTCTAACGGACACACAAATCAGA 3975
QY 2099 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGACAGCGGAGCGAGGTGAACATCG 2158
Db 3976 AGACTGAGTTACAAGCAATTCATCTAGCTTTGCGAGGATTCGGGATTAAGTAAACATAG 4035
QY 2159 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCCGACAGAGCGGAGCGG 2218
Db 4036 TGACAGACTCAATATGATTTGGAAATCATTTCAAGACCAACAGATAGAGTGAATCAG 4095
QY 2219 AGCTGGTGAACCCAGATCATCGAGAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGG 2278
Db 4096 AGTTAGTCAGTCAATAATAGAGCAGTTAATAAAAAAGAAAAAGTCTACCTGGCATGGG 4155
QY 2279 TGCCCGCCCAAGGGCATCGCGGCAAGCAGCAGATCGACAGCTGGTGTAGCAAGGGCA 2338
Db 4156 TACCAGCACAAAAGGAATTTGGAGAAATGAACAAATAGATGGTGGTTCAGTGTGGAA 4215
QY 2339 TCCGCAAGGTGCTGTTCTCTGGACGGCATCGAT 2370
Db 4216 TCAGGAAGTACTATTTTAGATGAATAGAT 4247

RESULT 10

US-08-935-312-13
; Sequence 13, Application US/08935312
; Patent No. 6207455
; GENERAL INFORMATION:
; APPLICANT: CHANG, Lung-Ji

;; TITLE OF INVENTION: LENTIVIRAL VECTORS
;; NUMBER OF SEQUENCES: 20
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BROWDY AND NETMARK, P.L.L.C.
;; STREET: 624 Ninth Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20001
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA: US/08/935,312
;; APPLICATION NUMBER: US/08/935,312
;; FILING DATE: 22-SEP-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: COOPER, Iver P.
;; REGISTRATION NUMBER: 28,005
;; REFERENCE/DOCKET NUMBER: CHANG=112
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 12494 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; DESCRIPTION: /desc = "DNA"
;; US-08-935-312-13

Query Match 45.6%; Score 1125.6; DB 4; Length 12494;
Best Local Similarity 68.4%; Pred. No. 1.1e-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

Qy 14 TGGCCGAGGCCATGAGCCAGGCGCACCA--CGGCCACATCTCTGATCAGCGCAGCAACT 70
Db 1548 TGGCTGAAGCAATGAGCGCAAGTAACAATCCAGCTACCATANTGATACAGAAAGCGCAATT 1607

Qy 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGGAGGCGGCACATCGGCC 130
Db 1608 TTAGAACCAAGAAAGACTGTTAAGTGTTCATTGTTGGCAAGAGAGGGGCACATAGCCA 1667

Qy 131 GCAACTGCGGCGCCCCCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGAGGGCGCCACGAGA 190
Db 1668 AAAATTGCGGGGCCCCTAGGAAAAAGGCTGTTGGAATGTGGAAGAGGAGGACACCAAA 1727

Qy 191 TGAAGGACTGACCGAGGCGCCAGCCCACTTCTTCCGCGAGGACCTGGCTTCCCGCAGG 250
Db 1728 TGAAGATTTGACTGAGAGACAGGCTAA-TTTTITAGGGAAGACTTGGGCTTCCCAAG 1786

Qy 251 GCAAGGCGCGGAGTTCGCCAGCAGCAGAACCGCGCAACAGCCGCCACCGCGCGAGC 310
Db 1787 GGAAGGCCAGGGAATTTCTTTCAGAGCAGACGAGCCACAGCCCCCAGAGAGAGAGC 1846

Qy 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGGCGCGGCGCGCGAGCGCCAGGCA 364
Db 1847 TTCAGGTTTGGGGAAGAGACAACTCTCTCAGAGAGCAGGAGCGCATAGACAAAGGAA 1906

Qy 365 -----CCTCAACTTCCCGCAGATCACCCGTGCGACGCGCCCTGGTGAGCATCAAGG 418
Db 1907 CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCGACCCCTCGTCAATAAAGA 1966

Qy 419 TGGCGGCGCAGATCAAGAGGCGCCCTGCTGGACACCGCGCGCGCGACACACCTGCTGGAGG 478
Db 1967 TAGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACATATTAGAAG 2026

Qy 479 AGATGAGCCTGCCGGGCAAGTGGAGGCCAAGATGATCGCGGGCATCGCGGCTTCATCA 538
Db 3107 CAGAAGACGAGCTAGAACTGGCAGAAACAGGAGGATTTCTAAAGNACCGGTACATG 3166

Db 2027 AAATGAATTTGCCAGGAAGATGGAACCAAAAAATGATAGGGGAATTTGGAGTTTATCA 2086

Qy 539 AGTGGCCCGCAGTACGACCATCTGATCGAGATCTCGGCAAGAGCCCATCGGACCG 598
Db 2087 AGTGAAGACAGTATGATCAGATCTCATGAAATCTCGGACATAAAGCTATAGGTACAG 2146

Qy 599 TGTGTGATGGGCCCCCAGCCCGTGAACATCATCGGCCCAACATCTGCTGACTCAGATTGGCT 658
Db 2147 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2206

Qy 659 GCACCTGAACTTCCCATCAGCCCCATCGAGCCCGTGGCGGTGAAGCTGAAGCCCGCA 718
Db 2207 GCACCTTAAATTTTCCCATAGTCTTATGAGACTGTACCAAGTAAATTAAGCCAGGAA 2266

Qy 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGGAGGAGAGATCAAGGCCCTGACCG 778
Db 2267 TGGATGGCCCAAAAGTTAAACAATGGCCATTGACAGAGAGAAAAAATAAGCATTAGTAG 2326

Qy 779 CCATCTCGAGGAGATGGAGAAAGAGGCGCAAGATCACCAGATCGGCCCGCGAGAACCCCT 838
Db 2327 AAATTTGTACAGAAATGGAAGAGGAGGAAAAATTTCAAAATTTGGCCTGAAAATCCAT 2386

Qy 839 ACAACACCCCGTGTTCGCCATCAAGAAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
Db 2387 ACAATACTCCAGTATTTGCCATAAAGAAAAAGACAGTACTTAAATGGAGAAAAATTAGTAG 2446

Qy 899 ACTTCGCGGAGCTGAACAAGCGCACCGAGGACTTCTGGAGGTGCAGCTGGGCGCATCCCC 958
Db 2447 ATTTACAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCAC 2506

Qy 959 ACCCGCGCGGCTTGAAGAAGAAAGAGCGGTGACCGTCTGGAGCTGGGCGCACCGCTACT 1018
Db 2507 ATCTGCGAGGTTTAAACACAGAAAAAATCAGTAACAGTACTGGATGTGGCGCATGCATATT 2566

Qy 1019 TCACGCTGCCCTGGACGAGGACTTCGCCAAGTACACCGCCCTTACCATCCCCAGCATCA 1078
Db 2567 TTTTCAGTTCCTTAGATAAAGACTTCAGGAAGTATCTGCACTTTACCATCTAGTATAA 2626

Qy 1079 ACAACGAGACCCCGGATCGCTACCACTACAACGTCTGCTCCCGCAGGCTGGAGGGCA 1138
Db 2627 ACAATGAGACACAGGAGTATAGATATCAGTACATGTTCTCCACAGGATGGAAGGAT 2686

Qy 1139 GCGCCAGCATTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCGCAACC 1198
Db 2687 CACCAAGCAATTTCCAGTGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAACAATC 2746

Qy 1199 CCGAGATCGTATCTACCAGTACATGGACGACCTGTACTGGCGACGACCTGGAGATCG 1258
Db 2747 CAGACATAGTCTATCTAATACATGGATGATTTGTATGATGATCTGACTTAGAATAG 2806

Qy 1259 GCCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGGCTTCAACA 1318
Db 2807 GCGAGCATAGACAAATAAGGNACTGACACATCTGTTGAGGTGGGATTTTACCA 2866

Qy 1319 CCGCCGACAAAGACACAGAGGAGCCCTTCTCTGTGATGGCTTACGAGCTGCAC 1378
Db 2867 CACCAAGCAAAAAACATCAGAAAGAACCTCCATCTCTTTGGATGGGTATTGAACCTCCATC 2926

Qy 1379 CCGACAGTGCACCTGCGAGCCCATCGAGCTGCGCGAGAGGAGAGCTGGACCGTGAAGC 1438
Db 2927 CTGATAAATGACAGTACAGCTATAGTCTGCCAAGAAAGACAGCTGGAGCTGCAATG 2986

Qy 1439 ACATCCAGAACTGTTGGGCAAGCTGAATCGGCCAGCAGATCTTACCCCGCATCAAGG 1498
Db 2987 ACATACAGAAATTAGTGGGAAAAATTTGAATTTGGCAAGTCAGATTTATGACGGGATTAAG 3046

Qy 1499 TGGCGCAGCTGTGAAGCTGTCGCGCGCGCAAGGCCCTGACCCACATCTGTCGCCCTCA 1558
Db 3047 TAAGGCAATTATGTAAGCTTCTTAGGGGAACCAAAAGCACTAACAGAAAGTAGTACCACATA 3106

Qy 1559 CCGAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGCGCGAGCGCGGCTGCAGC 1618
Db 3107 CAGAAGACGAGCTAGAACTGGCAGAAACAGGAGGATTTCTAAAGNACCGGTACATG 3166

```
QY 1619 GCGTGTACTACGACCCAGCAGACCTGGCGCGAGATCCAGACGAGCGCCACGACC 1678
Db 3167 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAGCGGGCAAGGCC 3226
QY 1679 AGTGACCTACAGATCTACAGGAGCGCTTCAAGAACTCTGAAGACCGGCAAGTACGCCA 1738
Db 3227 AATGACATATCAAAATTTATCAAGGCACTTTAAATCTGAAAACAGGAAATATGCA 3286
QY 1739 AGATGCGCACCGCCACACACAGCAGTGAAGAGCTGACCGCGCGCTGCGAGAGTCG 1798
Db 3287 GAATGAAGGGTGGCCACACTAATGATGTGAACAAATTAACAGAGCGAGTACAAAAATAG 3346
QY 1799 CCATGGAGAGCATGCTGATCTGGGCGCAAGCCCAAGTTCCTGCGCTGCGCCATCAGAGG 1858
Db 3347 CCACAGAAAGCATAGTAAATATGGGGAAGACTCTTAATTTAAATATCCCATCAAAAGG 3406
QY 1859 AGACTGGAGACCTGGTGGACGACTACTGGCAGGCGCACTGGATCCCCAGTGGAGT 1918
Db 3407 AAACATGGGAAGCATGGTGGACAGAGTATGGCAAGCCACCTGGATTCTGAGTGGAGT 3466
QY 1919 TCSTGAACACCCCGCCCTGGTGAAGCTGTGTACCAAGCTGGAGAGGAGCCCATATCG 1978
Db 3467 TTGTCAATACCCCTCCCTTTAGTGAAGTATGGTACCAAGTTAGAGAAAGAACCCATAATAG 3526
QY 1979 GCGCGGAGACCTTCTACGTGGACGCGCGCGCAACCGCGAGACCAAGATCGGCAAGCGC 2038
Db 3527 GAGCAGAACTTCTATGTAGATGGGCGAGCAATAGGGAACCTAAATAGGAAGAGCAG 3586
QY 2039 GCTACGTACGACGCGCGCGCGGCGGAGAGATCGTGAAGCTTACCGGAGACCAACACAGA 2098
Db 3587 GATATGTAACGTACAGAGGAGAGCAAAAAGTTGTCCCGCTTAACGGACACAAATCAGA 3646
QY 2099 AGACCGAGCTGCAGGCGCATCCAGTGGCGGCTGCGAGGAGCGCGGAGGAGTGAACATCG 2158
Db 3647 AGACTGAGTTACAAGCAATTCATCTAGCTTTGCGAGATTCGGGATTTAGAAGTAACATAG 3706
QY 2159 TGACCGAGCAGCAGTAGCGCTGGGCGCATCTCCAGGCGCGCGCGCAAGAGCGAGAGCG 2218
Db 3707 TGACAGACTCACAATATGATCTGGGATTCATTCAGCAGCAACAGATGAAGTGAATCAG 3766
QY 2219 AGCTGGTGAACAGCATCATGAGCAGCTGTATCAAGAGGAGAGGAGTGTACTGAGCTGGG 2278
Db 3767 AGTTAGTCAGTCAAAATATAGAGCAGTAAATTAAGGAGGAGGAGGAGTGTACTGCGATGG 3826
QY 2279 TGCGCGCCCAAGGCGCATCGCGGCAACGAGCAGATCGACAGCTGTGAGCAAGGSCA 2338
Db 3827 TACCAGCACAAAGGATTTGGAGGAATGAACAGTAGATGGTGGTTCAGTCTCTGGA 3886
QY 2339 TCCGCAAGGTGCTTCTCTGGACGGCATCGAT 2370
Db 3887 TCAGGAAGTACTATTTTAGATGGAATAGAT 3918
```

RESULT 11

```
US-08-848-760B-33
; Sequence 33, Application US/08848760B
; Patent No. 6248721
```

GENERAL INFORMATION:

```
; APPLICANT: Chang, Lung-Ji
; TITLE OF INVENTION: Animal Model For Evaluation Of Vaccines
```

```
; NUMBER OF SEQUENCES: 33
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
```

```
; STREET: 2421 N.W. 41st Street, Suite A-1
```

```
; CITY: Gainesville
```

```
; STATE: Florida
```

```
; COUNTRY: United States of America
```

```
; ZIP: 32606
```

```
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/848,760B
```

```
; FILING DATE: 25-Jan-2001
```

```
; CLASSIFICATION: <Unknown>
```

```
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 08/838,702
```

```
; FILING DATE: 09-APR-1997
```

```
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: PACE, DORAN R.
```

```
; REGISTRATION NUMBER: 38,261
```

```
; REFERENCE/DOCKET NUMBER: CNG-100C1
```

```
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (352) 375-8100
```

```
; TELEFAX: (352) 372-5800
```

```
; INFORMATION FOR SEQ ID NO: 33:
```

```
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 12494 base pairs
```

```
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
```

```
; TOPOLOGY: circular
```

```
; MOLECULE TYPE: other nucleic acid
```

```
; DESCRIPTION: /desc = "DNA"
```

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:
```

```
US-08-848-760B-33
```

```
Query Match 45.6%; Score 1125.6; DB 4; Length 12494;
```

```
Best Local Similarity 68.4%; Pred. No. 1,1e-176;
```

```
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;
```

```
QY 14 TGGCCGAGGCGCATGAGCCAGGCCACCA---CGGCCAACATCTGTGATGAGCGCAGCAACT 70
```

```
Db 1548 TGGCTGAAGCAATGAGCCAGTAACAATCAGCTACCATATGATACAGAAAGCAATT 1607
```

```
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTTCACTGCGCAAGAGAGGCGGCACATCGGCC 130
```

```
Db 1608 TTAGGAACCAAGAAAGACTGTTAAGTCTTCAATTTGGCAAGAGGCGCACATAGCCA 1667
```

```
QY 131 GCAACTGCCGGCCCCCGCGCAAGAGGCTCTGTGAAGTGGCGCAAGAGGCGCACACAGA 190
```

```
Db 1668 AAAATTGCGAGGCCCCCTAGGAAAGAGGCTCTTGAATGTGGAAGAGGAGCACACAAA 1727
```

```
QY 191 TGAAGGAGTGCACCGAGCGCCGACCACTTCTCCGCGAGGACCTGCGCTTCCCGCAGG 250
```

```
Db 1728 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTGTAGGGAAGATCTGCGCTTCCACAG 1786
```

```
QY 251 GCAAGGCGCGGAGTTCCTCCAGCGAGCAGAGAACCGGCCAACAGCCCGCCAGCGCAGC 310
```

```
Db 1787 GGAAGGCGAGGGAATTTCTTCAGAGCAGACAGCCACAGCCCGCCAGAGAGAGC 1846
```

```
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGGAGGCGCGCGCGCGCGCGCGCA 364
```

```
Db 1847 TTCAGGTTGGGGAAGAGACAACTCCCTCTCAGAAAGAGGAGCGCGATAGACAAGGAA 1906
```

```
QY 365 -----CCCTGAACCTTCCCGCAGATCACCTGTGCGAGCGCCCTGCTGTGAGCATCAAG 418
```

```
Db 1907 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGCGAGCGACCCCTGCTCAATAAAGA 1966
```

```
QY 419 TGGCGCGCGAGATCAAGAGGCGCTCTGTGGACACCGCGCGCGAGACACCGCTGTGGAG 478
```

```
Db 1967 TAGGGGGAATTAAGGAAGCTCTATTAGATACAGGAGCAGATGATACATATTAGAAG 2026
```

```
QY 479 AGATGAGCCTGCCCGCAAGTGGGAAGCCCAAGATGATCGCGGGGATCGGCGGCTTCATCA 538
```

```
Db 2027 AAATGAATTTGCCAGGAAGATGGAAACCCAAATATGATGGGGGAATTTGAGGTTTATCA 2086
```

```
QY 539 AGGTGCGCCAGTACGACAGATCTCTGATCGAGATCTCGGCAAGAGGCCATCGGCACCG 598
```

```
Db 2087 AAGTAGACAGATGATGATCACTACTATGAAATCTCGGAGACATTAAGCTATAGTACAG 2146
```

```
QY 599 TGTGTATCGGCGCCACCCCGTGAACATCATTCGCCCGCAACATCTGACCCAGCTGGGCT 658
```

```
Db 2147 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2206
```


ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 015280-249000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:

LENGTH: 1581 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: -

LOCATION: 1..1581

OTHER INFORMATION: /note= "pNlnSG11"

US-08-646-538-35

Query Match 45.6%; Score 1125.6; DB 3; Length 1581;
Best Local Similarity 68.4%; Pred. No. 1..le-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

QY 14 TGGCCGAGCCATGAGCCAGGCCACCA--GCGCCAACATCCTGATGAGCGCAGCAACT 70
DB 1877 TGGCTGGAAGCAATGAGCCAAAGTAACAAATCCAGCTACCATAATGATACAGAAAGGCAATT 1936

QY 71 TCAAGGGCCCCAAGCGCATCACTAGTCTCACTGCGGCAAGAGGGCCACATCGGCC 130
DB 1937 TTAGGAACCAAGAAAGACTGTGTAAGTGTTCAAATTTGGCAAGAAAGGGCACATAGCCA 1996

QY 131 GCAACTGCGGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGAGGGCCACCCAGA 190
DB 1997 AAAATTGACAGGCCCTAGGAAAAGGCTGTGGAATGTGGAAGAGGACACACAAA 2056

QY 191 TGAAGACTGCACGAGCGCCAGGCCCACTTCTTCGCGAGGACCTGCCCTTCCGCCAGG 250
DB 2057 TGAAGATTGACTGAGAGACAGGGCTAA--TTTTTTAGGGAAGACTGGCCCTCCACCAAG 2115

QY 251 GCAAGGCCGCGGAPTCCCGAGGAGCAGAACCGCGCCCAAGAGCCCAACAGCCCAAGGAGC 310
DB 2116 GGAAGCCAGGGAATTTCTCAGAGCAGACAGAGCCCAAGCCCAAGAGGAGC 2175

QY 311 TGCAGGTGCGGG-----CGAACACCCCGCAGCGAGCGCGCGCCGCGCCAGGCCA 364
DB 2176 TTCAGGTTGGGAAGAGACAACTCCCTCTCAGAGCAGGAGCCGCTAGACAGGNA 2235

QY 365 -----CCCTGAACCTTCCCGAGATCACTCTGTCGAGCGGCCCTGTTGAGCATCAAGG 418
DB 2236 CTGTATCCTTTAGCTTCCTCAGATCACTCTTTGGCAGCGAGCCCTCGTCACAATANAGA 2295

QY 419 TGGCGGCGCAGATCAAGAGGCCCTGCTGGACACCGCGCGCCGACACACCTGCTGGAGG 478
DB 2296 TAGGGGGCAATTAAGGAAGCTCTTAGATACAGGAGCAGATACAGATTTAGAG 2355

QY 479 AGATGAGCTGCGCGCAAGTGGAAAGCCCAAGATGATCGGGGCACTCGCGGCTTCATCA 538
DB 2356 AAATGAATTTGCCAGAGATGGAACCAAAATGATAGGGGAATTTGAGGTTTATCA 2415

QY 539 AGGTGCGCCAGTACGACAGATCTCTGATGAGAGTCGCGGCAAGAGGCCATCGGCAACG 598
DB 2416 AAGTAGGACATGATGATCAGATCTCATAGAAATCTCGGGACATAAAGCTATAGGTACAG 2475

QY 599 TGCTGATCGGCCCAACCCCGTGRACATCATCGGCCGCAACATGCTGACCCAGCTGGGCT 658
DB 2476 TATTAGTAGGACCTACACTGCTCAACATAATTTGGAAGAAATCTGTTGACTCAGATTGGCT 2535

QY 659 GCACCCCTGAACCTTCCCATCATCGCCCATCTGAGACCGTGCCTGTAAGCTGAAGCCCGCA 718
DB 2536 GCACCTTAATTTCCCATATTAGTCTATTGAGACTGTACCGAGTAATTAAGCCAGGAA 2595

QY 719 TGGAGGCGCCCAAGGTGAAGCAGTGGGCCCTGACCGAGGAGAAATCAAGGCCCTGACCG 778

DB 2596 TGGATGGCCCAAAAGTTAAACAATGSCCATTGACAGAAAGAAAATAAAGCATTTAGTAG 2655
QY 779 CCATCTGCGAGGAGATGAGAGAGGGGGAAGATCACCAGATCGGCCCGGAGAACCCCT 838
DB 2656 AAAATTTGTACAGAAATGGAAGGAGGAAAATTTCAAAAATTTGGCCCTGAAAATCCAT 2715
QY 839 ACAACACCCCGCTGTTTCGCCATCAAGAAGAAGACAGACCAAGTGCGCAAGCTCGTGG 898
DB 2716 ACAATCTCCAGTATTGCCATAAAGAAAAGACAGTACTAATAGGAGAAATTTAGTAG 2775
QY 899 ACTTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGACAGCTGGGCAATCCCC 958
DB 2776 ATTTCCAGAGAACTTAATAAGAGAACTCAAGATTCTTGGGAAGTTCAATTAGGAATACCAC 2835
QY 959 ACCCGCGCGGCTGAAGAAGAAGAGAGCGTGACCGTGTGACGTGGGCGGAGCGCTACT 1018
DB 2836 ATCTTCGAGGGTTAAACACAGAAAAATCAGTAAACAGTACTGGATGTGGCGATGCATATT 2895
QY 1019 TCAGCGTGGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 1078
DB 2896 TTTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATACTGCATTTACCATACCTTAGTATAA 2955
QY 1079 ACAACGAGACCCCGGCACTCCGCTACCAAGTACAAAGCTGTGCCCGCAGGCTGGAAGGCA 1138
DB 2956 ACAATGAGACACAGGATTAGATATCAGTACAACTGCTTCCACAGGATGAAAGGAT 3015
QY 1139 GCCCGCAGATCTTCAGAGCAGCATGACCAAGATCTGAGGCCCTTCCGCGGCCGCAACC 1198
DB 3016 CACCAAGCATATTTCAGTGTAGTATGACAAAAATCTTAGAGCCCTTTAGAAAAACAATC 3075
QY 1199 CCGAGATCTGATCTACCAAGTACATGGAGAGCTGTGACGTGGGCGACGACCTGGAGATCG 1258
DB 3076 CAGACATAGTCACTATCAATACATGGATGATTTGTATGATGATCTGACTTAGAATAG 3135
QY 1259 GCCAGCACCGCCCAAGATCGAGGAGCTGCGAAGCACCTGTGCTGCGCTGGGCTTCCACCA 1318
DB 3136 GGCAGCATAGAACAAAAATAGAGAACTGAGCAACATCTGTTGAGTGGGATTTACCA 3195
QY 1319 CCCCAGACAGAGACACAGAGGAGGCCCTTCTGCTGGATGGCTAGGAGCTGAC 1378
DB 3196 CACCAAGAAAAACATCAGAAAGAACCTTCCATTTCTGTTGGATGGGTTTGAACCTCCATC 3255
QY 1379 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGGAGAGCTGACCGTGAACG 1438
DB 3256 CTGATTAATGGACAGTACAGCCCTATAGTCTGCCAGAAAGACAGCTGACTGTCAATG 3315
QY 1439 ACATCCAGAGCTGTGGGCAAGCTGAATGGGCGCCAGCCAGATCTACCCCGCATCAAG 1498
DB 3316 ACATACAGAAATTAGTGGGAAAATTTGAATGGCAAGTCAGATTTATGCAGGGATTAAAG 3375
QY 1499 TCGCGCAGCTGTGCAAGCTGTGCGGGGCGCCAGGCCCTGACCGACATCTGCGCCCTGA 1558
DB 3376 TAAGCAATATTGTAACCTTCTTAGGGGAAACCAAGCACTAACAGAAAGTAGTACCATAA 3435
QY 1559 CCGAGAGGCGCAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGCCGCTGCACG 1618
DB 3436 CAGAGAGCAGAGCTAGACTGGCAGAAACAGGAGATTTCTAAAGAACCGGTACATG 3495
QY 1619 GCGTACTTACGACCCCAAGAGGACCTGGTGGCGGAGATCCAGAGACAGGCGCCAGCC 1678
DB 3496 GAGTGTATTATCACCCCTCAAAAGACTTAATAGCAGAAATACAGAGCAGGGGCAAGGCC 3555
QY 1679 AGTGACCTACAGATCTACAGGAGGCCCTTCAAGAACCTGAAGACCGGCAAGTAGGCCA 1738
DB 3556 AATGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAAATATGCAA 3615
QY 1739 AGATCGCAGCCCGCACACCAAGCAGTGAACAGCTGACCGAGGCCGTCGAGAGATCG 1798
DB 3616 GAATGAAGGTCGCCACACTATGATGTGAAACAAATTAACAGAGGCGAGTACAAAAATAG 3675
QY 1799 CCATGAGAGCATCTGATCTGGGCAAGACCCCAAGTTCCGCCCTGCCCATCCAGAGG 1858
DB 1858

```
Db 3676 CCACAGAAAGCATAGTAAATATGGGAAAGACTCTAAATTTAAATTACCCATACAAAAGG 3735
;
; REFERENCE/DOCKET NUMBER: 015280-249000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..15581
; OTHER INFORMATION: /note= "pNlnSG11"
; US-09-503-222-35

;
; Query Match 45.6%; Score 1125.6; DB 4; Length 15581;
; Best Local Similarity 68.4%; Pred. No. 1.le-176;
; Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

QY 14 TGGCCGAGCCATGACGAGCCACCA---GCGCAACATCTCTGATCGAGCGCAGCAACT 70
Db 1877 TGGCTGAGCAATGACCCCAAGTAACAAATCCAGTACCATTANTGATACAGAAAGCAATT 1936
QY 71 TCAAGGCCCCCAAGCGCATCATCAAGTGTCTTCAACTGCGCAAGGAGGGCCACATCGCCC 130
Db 1937 TTAGGAACCAAGAAAGAGACTGTTAAGTGTTCATTTGTCGCAAGAGGGCAGCATAGCCA 1996
QY 131 GCAACTGCCGCGCCCCCGCCCAAGAGGGCTGCTGGAAGTGGCGCAAGGGGCCACCAGA 190
Db 1997 AAAATTTGCAAGGGCCCCCTAGGAAAAAGGGCTGTTGGAATTTGGAAGAGGAGGACAAA 2056
QY 191 TGAAGACTGCAAGCGGCGGCGGCAACTTCTTCCGAGGAGCCTGGCTTCCCGCAGG 250
Db 2057 TGAAGATTGTACTGAGACAGGCTAA-TTTTGTAGGGAAGATCTGGCTTCCACAAG 2115
QY 251 GCAAGCGCCGAGTTTCCCGCAGGAGCAACCGCGCCAAAGAGCCCCACAGCGCCGAGC 310
Db 2116 GGAAGCGCAGGGAATTTTCTTCAGAGCAGACACAGAGCCCAAGCCCCCAGCAAGAGAGC 2175
QY 311 TCGAGTGGCGGG-----CGACAACCCCGCAGGAGCGCGCGCGGCGGAGCCAGGCA 364
Db 2176 TTCAGGTTTGGGGAAGAGACAACACTCCCTCTCAGAAAGCAGGAGCGATAGACAAGAA 2235
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTGTGAGCATCAAG 418
Db 2236 CTGTATCTTTCCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACATAAGA 2295
QY 419 TGGGCGCCGAGATCAAGGAGGCGCTGTGGACACCGCGCGCGGAGCACCGTGTGGAGG 478
Db 2296 TAGGGGGCAATTTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 2355
QY 479 AGATGAGCCTGCCCCGCAAGTGAAGCCCCAAGATGATCGCGGCATCGCGGCTTCATCA 538
Db 2356 AAATGAATTTTGGCAGGAAGATGGAACCAAAATGATAGGGGGAATTTGGAGGTTTATCA 2415
QY 539 AGTGCGCCAGTACGACGAGATCCTGATCGAGATCTCGGCAAGAGGCCATCGCGACCG 598
Db 2416 AAGTAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2475
QY 599 TGTGATCGGCCCCCGCGCTGAACATCATTCGCGCCCAACATGCTGACCCAGCTGGGCT 658
Db 2476 TATTAGTAGGACCTACACCTGTCAACATAATTTGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGAACTTCCCATCATCGCCCATCGAGACCGTGGCGCGCTGAAGTGAAGCCCGCA 718
Db 2536 GCACCTTAAATTTTCCCATTTAGTCTTATGAGATGTGACAGAGGAGGAGGAGGAGGAGGAG 2595
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCCCTGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 778
Db 2596 TGGATGCGCCCAAGGTGAAGCAGTGGCCCCCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2655
```

```
RESULT 13
US-09-503-222-35
; Sequence 35: Application US/09503222
; Patent No. 6265548
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Gaitanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
; TITLE OF SEQUENCES: 37
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,222
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,538
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677
```

```
Qy 779 CCATCTCGAGGAGATGGAGAGGAGGAGATCACCAGATCGGCCCGCCGAGAACCCCT 838
Db 2656 AAATTTGTACAGAAATGAAAGGAGGAAATTTCAAAAATTTGGGCTGAAATCCAT 2715
Qy 839 ACAACACCCCGTGTTCGCCATCAAGAGAGAGAGACAGCACCAGTGGCGCAGCTGGTGG 898
Db 2716 ACAATACTTCCAGTATTTGCCATTAAGAAAGAAAGACAGTACTAAATGGAGAGAAATTAGTAG 2775
Qy 899 ACTTCGGGAGCTGAACAAGCGCCACAGGACTTCTGGGAGGTGCAGCTGGGCTATCCCC 958
Db 2776 ATTCAGAGAACTTAATAGAGAACTCAAGATTTCTGGGAAATTCATATTAGAAATACCAC 2835
Qy 959 ACCCGCGCGGCTTGAAGAAGAGAGAGCGTGACCGTGTCTGGACGTGGCGACGCTACT 1018
Db 2836 ATCTGCAAGGTTTAAACAGAGAAAAATCAGTAACAGTACTGGATGTGGCGCATGATATT 2895
Qy 1019 TCAGGTGCCCTGGACGAGAGACTTCCGCAAGTACACCGCTTACCAATCCCCAGCATCA 1078
Db 2896 TTTCAAGTTCCCTTAGATAAAGACTTCAGAAAGTATCTGCAATTTACCATACTAGTATAA 2955
Qy 1079 ACAACGAGACCCCGGATCCGCTACCAAGTACAAGCTGCTGCCCGAGGCTGGAGGSCA 1138
Db 2956 ACATGAGACACAGGATTAAGATATCAGTACAAATGTCTTCCACAGGATGGAAAGAT 3015
Qy 1139 GCGGAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCGCAACC 1198
Db 3016 CACGAGCAATATCCAGTGTAGCATGACAAAAATCTTAGAGCCTTTATAGAAAAACAATC 3075
Qy 1199 CCGAGATCGTATCTACCAGTACATGGAGACCTGTAGCTGGGACGACGCTGGAGATCG 1258
Db 3076 CAGACATAGTCATCTCAATACATGAGTATTTGTATGTAGGATCTGACTTACGAAATAG 3135
Qy 1259 CCCAGACCGCGCAAGATCGAGAGCTGCGCAAGACCTCTGCGCTGGGCTTCACCA 1318
Db 3136 GCGAGCATAGAACAAAATAGAGAACTGACACACATCTCTTGAGGTGGGATTTACCA 3195
Qy 1319 CCGGACAGAGAACACAGAGAGGAGGCGCCCTTCTGTGGATGGGCTACGAGCTGACCC 1378
Db 3196 CACGAGACAAAACATCAGAAAGAACCTCCATCTCTTTGGATGGTTTATGAACCTCCATC 3255
Qy 1379 CCGACAGTGCACGCTGCAGCCCATCGAGCTGCGGAGAGAGAGAGCTGGACGTGAACG 1438
Db 3256 CTGATAATGACAGTACAGCCTTAGTGTGCGCAGAAAGAGACGCTGGACTGTCAATG 3315
Qy 1439 ACATCCAGACCTGGTGGGCAAGCTGAACCTGGCGAGCAGATCTACCCGCGCATCAAG 1498
Db 3316 ACATACAGAAATTAGTGGGAAATTAATTTGGCAAGTCAATTTATCGAGGATTAAG 3375
Qy 1499 TCGCCAGCTGTGAAGCTGTGCGCGCGCCAGGCGCTGACGACATCTGTCGCCCTGA 1558
Db 3376 TAAGCAATTTATGAACTTCTTAGGGGACCAAGACATCAACAGAGTAGTACCCTAA 3435
Qy 1559 CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGGAGCGCCGTGCACG 1618
Db 3436 CAGAAGAGCAGAGCTGAGACTGCGAGAAAACAGGGAGATTCTAAAAGAACCGGTACATG 3495
Qy 1619 CCGTGTACTACGCCACCGAGGAGCTGTGGCGGAGATCCAGAAAGCAGGCGCACGACC 1678
Db 3496 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAGCAGGGGCAAGGCC 3555
Qy 1679 AGTGGACCTACAGATCTACAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1738
Db 3556 AATGGACATATCAAAATTTATCAAGGCACTTTAAAATCTCAAAACAGAAAATATGCA 3615
Qy 1739 AGATGGCAGCGCCACACAGCAGTGAAGAGCTGACCGCGCGCTGGCAGAGATCG 1798
Db 3616 GAATGAAGGGTGGCCACACTAATGATGTGAACCAATTAACAGAGCGAGTACAAAAATAG 3675
Qy 1799 CCATGGAGAGCATCGTGTATCTGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAGG 1858
Db 3676 CCACAGAAAGCATAGTAAATATGGGAAAGACTCTCTAAATTTAAATACCCATACAAAG 3735
Qy 1859 AGACCTGGAGACCTGGTGGAGCCGACTACTGGCAGGCGCACCTGGATCCCCAGTGGGAGT 1918
```

```
Db 3736 AAACATGGGAAGCATGGTGGACAGAGTATTCGCAAGCCACCTGGATTCTCGAGTGGAGT 3795
Qy 1919 TCGTGAACACCCCGCTCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGAGCCCATCATCG 1978
Db 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGTACAGTTAGAGAAAGAACCCATANTAG 3855
Qy 1979 GCGCGAGACCTTCTACGTGGACGCGCCCAACGCGGAGACCAAGATCGGCAAGGCG 2038
Db 3856 GAGCAGAACTTTCTATGTAGTGGGAGCCCAATAGGAAACTTAATTAGGAAAGCAG 3915
Qy 2039 GCTAGTGACCGACCGGGCGGCGCAGAAAGATCGTGAGCTTGACCGAGACCAACCCAGA 2098
Db 3916 GATATGTAACTGACAGAGAGACAAAAAGTTGTCCCTTAACGAGACACACAACATCAGA 3975
Qy 2099 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGAGACGCGGACGAGGTGAACATCG 2158
Db 3976 AGACTGAGTTACAAGCAATTCATCTAGCTTTGACAGGATTCGGGATTAAGATTAACATAG 4035
Qy 2159 TGACCGACGACGAGTACGCCCTGGGCATCCAGCGCCAGCCGACAGAGCGAGAGCG 2218
Db 4036 TGACAGACTCACAATATGCATTTGGAAATCATTCAGACACCAACCATAGAGTGAATCAG 4095
Qy 2219 AGCTGGTGAACCCAGATCATCGAGCAGCTCATCAAGAAGAGAGAGTGTACCTGAGCTGG 2278
Db 4096 AGTTAGTCAGTCAAAATAATAGAGCAGTTAAATAAAGAGAAAGTCTACCTGGCATGG 4155
Qy 2279 TGCCCGCCCAAGGCGATCGCGCGCAGAGCAGATCGCAAGCTGGTGAGCAAGGCA 2338
Db 4156 TACCAGCACACAAAGGAATGGAGAAATGAACAAGTAGAGTGGTGGTCTAGTGTGGAA 4215
Qy 2339 TCCGCAAGGTGCTGTCTCTGGACGCGCATCGAT 2370
Db 4216 TCAGGAAAGTACTATTTTAGATGAATAGAT 4247
```

RESULT 14
US-08-944-449-7
; Sequence 7, Application US/08944449
; Patent No. 5985613
; GENERAL INFORMATION:
; APPLICANT: KURTH, REINHARD
; APPLICANT: BAIER, MICHAEL
; APPLICANT: METZNER, KARIN
; APPLICANT: WERNER, ALBRECHT
; TITLE OF INVENTION: Use of an "immunodeficiency-virus suppressing
; TITLE OF INVENTION: lymphokine (ISL)" to inhibit the replication of
; TITLE OF INVENTION: viruses, particularly of retroviruses
; FILE REFERENCE: 8341-7065
; CURRENT APPLICATION NUMBER: US/08/944,449
; CURRENT FILING DATE: 1997-10-06
; EARLIER APPLICATION NUMBER: EP 95113013.2
; EARLIER FILING DATE: 1995-08-18
; EARLIER APPLICATION NUMBER: DE 195 13 152.5
; EARLIER FILING DATE: 1995-04-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-08-944-449-7

Query Match 45.2%; Score 1116; DB 2: Length 9737;
Best Local Similarity 68.1%; Pred. No. 4.le-175;
Matches 1616; Conservative 0; Mismatches 740; Indels 16; Gaps 4;

Qy 14 TGGCGAGCGCATGAGCCAGCCACCA--GGCCCAACATCTGATGAGCCGACCAACT 70
Db 1884 TGGCTGAAGCCATGAGCCAGTAACAAATCCAGCTAACATATGATGAGAGGCAATT 1943
Qy 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTCAACTGGCGCAAGGAGGCCCATCGCCC 130

Db 1944 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATTTGTGGCAAGAGGCAATAGCCA 2003
QY 131 GCAACTGCCCGCCGCCCAAGAGGGCTGCTGGAAGTCGGCAAGAGGGCCACACAGA 190
Db 2004 AAAATTGCAGGGCCCTAGGAAAAGGGCTGTTGGAGATGTGGAAGGAGACACCAA 2063
QY 191 TGAAGAGTCCACCGAGCGCCAGGCAACTTCTCCGCGAGGACCTGGCCCTTCCCGCAGG 250
Db 2064 TGAAGATTGCATGAGACAGAGCTAA-TTTTATTAGGAAAGATCTGGCCCTTCTACAG 2122
QY 251 GCAAGGCCCGAGTTCCCGAGCGAGCAGAACCGCCGACACGCCCCACCGCCGAGC 310
Db 2123 GGAAGGCCAGGAAATTTCTTCAGAGCAGACACAGAGCCGACAGAGAGAGC 2182
QY 311 TGCAGGTGCGCGCG-ACAAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCGCA 364
Db 2183 TTCAGGTTTGGGAGGAGAAACAACCTCCCTCTCAGAAGCAGGAGCGCATAGACAAGAA 2242
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTCTGGTGAGCATCAAGG 418
Db 2243 CTGTATCTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAAATAGGA 2302
QY 419 TGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGAGCACACCGTCTGGAGG 478
Db 2303 TAGGGGGCAACTAAGGAAGCTCTATTAGATACAGAGCAGATGATACAGTATTAGAAG 2362
QY 479 AGATGAGCCTGCCCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
Db 2363 AAATGAATTTGCCAGGAAATGAAACCAAAATGATAGGGGAATTTGAGGTTTTATCA 2422
QY 539 AGTGGCCAGTACGACAGATCCTGATCGAGATCTGCGCGCAAGAGGCCATCGGCACCG 598
Db 2423 AAGTAAGACAGTAGATCACTGTAGAAATCTGTGACATAAAGCTATAGGTACAG 2482
QY 599 TGCTGATCGSCCCACCCCGTCAACATCATCGCGCGCAACATGCTGACCCAGCTGGCT 658
Db 2483 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTGACTCAGATTGGTT 2542
QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGCOCGTGAAGCTGAAGCCCGGCA 718
Db 2543 GTACTTTAAATTTCCCATTTAGTCTATTGAAACTGTACCAGTAAAAATTAAGCCAGAA 2602
QY 719 TGACGGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCG 778
Db 2603 TGGATGGCCCAAGATTAAGCAATGGCCATTGACAGAAGAAAAATAAAGCATTAGTAG 2662
QY 779 CCATCTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGAGAACCCCT 838
Db 2663 AGATATGTACAGAAATGGAAGAGAGGGAATTTCAAAAATTTGGCCCTGAAATTCAT 2722
QY 839 ACAACACCCCGCTGTTCCGCATCAAGAAGAAAGACACACACCAAGTGGCGCAAGCTGGT 898
Db 2723 ACAATACTCCAGTATTTCCTATAAAGAAAAAGACAGTACTAAATGGAGAAAACTAGTAG 2782
QY 899 ACTTCGCGAGCTGAACAGCGCACCCAGGACTCTTGGGAGGTGACGCTGGGCATCCCC 958
Db 2783 ATTTCAGAGAACTTAATAAAGAACTCAAGACTTCTGGGAAGTTCAGTTAGGAATACCAC 2842
QY 959 ACCCCGCGGCTTGAAGAAGAACGAGCTGACCGTGTCTGGACGTGGCGGACCGCTACT 1018
Db 2843 ACCCCGAGGTTTAAAAAGAAAAATCAGTAACAGTATTGGATGTGGGTGATGCATACT 2902
QY 1019 TCAGCGTGGCCCTGGAGGAGCTTCCGAAGTACACCGCTTACCATCCCGCAGCATCA 1078
Db 2903 TTTTCAGTTCCTTAGATAAAGACTTTAGAAAGTATACATGATTTACCATACCTAGTATAA 2962
QY 1079 ACAACAGAGCCCCGATCCGCTACAGTACAAAGTGTGCCCCGAGGGCTGGAAGGGCA 1138
Db 2963 ACAATGAGACACACGGGATTTAGATATCAGTACAAATGTGCTGCCACAGGGATGGAAGGAT 3022
QY 1139 GCCCGAGCATTTCCAGAGCAGATGACCAAGATCTCTGGAGCCCTTCCCGCGCCGCAACC 1198
Db 3023 CACCAGCATATTTCCAAAGTAGCATGACAAAAATCTTAGAGCCCTTTTAGAAAAAGAAATC 3082

QY 1199 CCAGAGTCTGTGATCTACAGTACATGACGACCTGTAGCTGGCAGCGACCTGGAGATCG 1258
Db 3083 CAGACATAGTTATCTATCAATACATGGATGATTGTATGTAGGATCTGACTTAGAATAG 3142
QY 1259 GCAGACACCGCGCAAGATCGAGAGCTGCGCAAGCACTGCTGCGCTGGGGCTTACCA 1318
Db 3143 GGCAGCATAGAACAAAATAGAGAACTAGACAGCATCTGTTGAGGTGGGATTTACCA 3202
QY 1319 CCCCCGACAAGAGACACCAAGAGGCCCTTCCCTGTGATGGGTACGAGCTGCACC 1378
Db 3203 CACAGACAAAACATCAGAAAGACCTCCATCTCTTTGGATGGGTTATGAACCTCATC 3262
QY 1379 CCACAAAGTGGACCGTGCAGCCCATCGAGTGCCTCCGAGAGGAGAGCTGACCGTGAACG 1438
Db 3263 CTGATAAATGGACAGTACAGCTTAAATGCTGCCAGAAAAGACAGCTGGACTGTCAATG 3322
QY 1439 ACATCAGAAAGCTGTGGGCAAGCTGAACTGGGCCAGCAGATCTTACCCCGGATCAAGG 1498
Db 3323 ACATACAGAAAGTTAGTGGGAAAATTTGAAATTTGGCAAGTCAAGATTTATGCAAGGATTAAG 3382
QY 1499 TGGCCGAGCTGTCAAGCTGCTGCGGGGCCCAAGGCCCTGACCGACATCTGCGCCCTGA 1558
Db 3383 TAAAGCAGTTATGTAACTCCTTAGAGGAACCAAGCCTAACAGAAATTAATACCATAA 3442
QY 1559 CCAGGAGCGCGAGCTGGAGCTGGCCGAGAACCCGAGATCTCTGCGGAGCCCGTGCACG 1618
Db 3443 CAGAAGAACGACAGCTAGAACTGGCAGAAAACAGGGAGATCTTAAAGAAACACAGTACATG 3502
QY 1619 GCGTGTACTAGACCCCGCAGCAAGGACCTGTTGGCCGAGATCCAGAACGAGGCCACACC 1678
Db 3503 AAGTATATTATGACCCATCAAAAGACTTAGTAGCAAAATACAGAAGCAGGGGCAAGGCC 3562
QY 1679 AGTGGACCTACCAGATCTACCAGAGGCCCTTCAAGAACCTTGAAGACCGCAAGTACGCCA 1738
Db 3563 AATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAAACAGAAAGATGCA 3622
QY 1739 AGATGGCACACCCCAACACAGCTGAAGCAGCTGACCGAGGCGTGCAGAAAGATCG 1798
Db 3623 GGATAGGGGTGCCCCACACTAATGATGATAAACAGTTTAAACAGAGGAGTGCAGAAAGTAT 3682
QY 1799 CCATGAGAGCATCTGTTGGGCAAGACCCCAAGTTCCCGCTGCCCATCCAGAAAG 1858
Db 3683 CCACAGAAAGCATAGTAATATGGGAAAGATCTCTAAATTTAAACTACCCATCAAAAG 3742
QY 1859 AGACCTGGGAGACTGTGTGACCGACTACTTGGCAGGCGACCTGGATCCCGAGTGGAGT 1918
Db 3743 AAACATGGGAGCATGGTGGATGGAGTATTGGCAAGCTACTTGGATCTCTGAGTGGAGT 3802
QY 1919 TCGTGAACACCCCGCCCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGGCCCATCATCG 1978
Db 3803 TTGTCAATACCCCTCCCTTAGTGAAATTTATGTTACCAGTTAGAGAAAAGACCCATAGTAG 3862
QY 1979 GCGCCGAGACTTCTAGCTGGAGCGGCCCAACCCGAGACCAAGATCGGCAAGGCCG 2038
Db 3863 GAGCAGAACTTTCTATGTAGTAGTGGGCAAGCTAATAGGGAGACTTAAATTAGGAAAAGCAG 3922
QY 2039 GCTACGTGACCGACCGGGCGCGCAGAAAGATCGTAGCCTGACCGAGACCAACCAACAGA 2098
Db 3923 GATATGTACTGACAGGAGAGACAAAAGTTGTCTCCATAGCTGACACAAACAAATCAGA 3982
QY 2099 AGACCGAGCTGCAAGGCATCCAGCTGGCCCTGCAAGGACACGGCAGGAGGTGAACATCG 2158
Db 3983 AGACTGAAATTAACAAGAAATTCATCTAGCTTTGCAAGGATTCGGGATTAAGAAGTAAACATAG 4042
QY 2159 TGACCACAGCCAGTACCGCCCTGGGCATCATCCAGGCCAGCCCGCAGCGAGGAGCG 2218
Db 4043 TAAACAGCTACAAATATGCAATTAGGAATCATCAAGCACCAACCAAGATTAAGAGTAAATCAG 4102
QY 2219 AGCTGTTGAACAGATCATCGACAGCTGATCAAGAGGAGAGAGGTGTACTGTAGCTGGG 2278
Db 4103 AGTTAGTCAATTAATAGACAGCTTAATAAAAAAGGAAAGGCTTACTGTGCATGGG 4162

QY 1379 CCGACAAAGTGGACCGTGCAGCCCATGAGCTGCCGAGAAAGAGAGAGCTGGACCGTGAACG 1438
Db 3263 CTGNTAATGGACAGTACACCCCTATATGCTGCCAGAAAAGACAGCTGGACTGTCAATG 3322
QY 1439 ACATCCAGAAGCTGGTGGGCAAGCTGAATCTGGCCGACCCAGATCTACCCCGGCATCAAGG 1498
Db 3323 ACATACAGAAGTGTAGTGGGAAAATGAAATGGGCAAGTCAAGATTTATGCAAGGATTAAG 3382
QY 1499 TGGCCAGCTGTGCAGCTGCTGGCGGCCCAAGGCCCTGACCGACATCGTGCCTCTGA 1558
Db 3383 TAAAGCAATGTTATGTAACCTCTTAGAGGAACCAAGCACCTAACAGAAATATACCACTAA 3442
QY 1559 CCGAGGAGCGGAGCTGGAGCTGCCGAGAACCGCAGACTCCTCGCGAGCGCCGCTGCAACG 1618
Db 3443 CAGAAGACAGAGCTAGACTGCGAGAAAACAGGGAGATTTCTAAAGAACCAAGTACATG 3502
QY 1619 GCGTGTACTAGCACCCAGCAAGACCTGCTGGCCGAGATCCAGAAAGCAGGCGCACGACC 1678
Db 3503 AAGTATATTTAGCCCATCAAAAGACTTAGTAGCAGAAATACAGAAAGCAGGCGCAAGGCC 3562
QY 1679 AGTGGACCTACCAAGCTACCAAGAGCCCTTCAGACCTGAAGACCGGCAAGTACGCCA 1738
Db 3563 AATGGACATATCAAAATTTATCAGAGCCATTTAAAAATCTGAAAACAGGAAGTATGCAA 3622
QY 1739 AGATGCGCACCCGCCACCAACGACGTGAAGCAGCTGACCGAGCGCCCTGCAGAAAGATCG 1798
Db 3623 GGATGAGGGGTGCCACACTAATGATGTAACACAGTTAACAGAGGCAAGTGCAAAAGTAT 3682
QY 1799 CCATGGAGAGCATCGTGATCTGGGGCAAGACCCCAAGTTCCGCCCTGCCCATCCAGAAGG 1858
Db 3683 CCACAGAAAGCATAGTAATATGGGAAAGATTCCTAAATTTAAACTACCCATACAAAAGG 3742
QY 1859 AGACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACTGGATCCCGAGTGGGAGT 1918
Db 3743 AAACATGGGAAGCATGGTGATGGAGTATTTGGCAAGCTACCTGGATTCCTGGAGTGGGAGT 3802
QY 1919 TCGTGAACACCCCCCTCGTGAAGCTGTGTTACAGCTGGAGAGGAGGCCCATCATCG 1978
Db 3803 TTGTCATACCCCTCCCTTAGTGAATTTATGTTACCAATTTAGAGAAAGACCCATAGTAG 3862
QY 1979 GCGCCGAGACCTTCTAGCTGGAGCGGCGGCCCAACCGCCGAGACCAAGATCGGAAGCCG 2038
Db 3863 GAGCAGAAACTTTCTATGTAGTGGGACGCTAATAGGGAGACTAAATTTAGGAAAAGCAG 3922
QY 2039 GCTAGGTACCGGCGGCGGCGCAGAGATCGTGAGCCTGACCGAGACCACCAACCA 2098
Db 3923 GATATGTTCTGACAGAGGAAGCAAAAAGTTGTCTCATAGCTGACACAAACAAATCAGA 3982
QY 2099 AGACCGAGCTGCAGGGCATCCAGCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCG 2158
Db 3983 AGACTGAATTAACAGCAATTCATCTAGCTTTCAGGATTCGGGATTAGAAGTAAACATAG 4042
QY 2159 TGACCGACAGCCAGTAGCCCTGGGCATCATCCAGGCCCGCCGACAGAGAGGAGAGCG 2218
Db 4043 TAACAGACTCACATATGCTATAGGAATCATTCGAACACCAACAGATAAGAGTGAATCAG 4102
QY 2219 AGCTGTGACCAAGATCATCGACGCTGATCAAGRAGGAGAGGTGTACTGAGCTGGG 2278
Db 4103 AGTTAGTCACTCAATTAAGACAGTTAATAAAAAGGAAAAGGCTTACCTGGCATGGG 4162
QY 2279 TGCCCGCCCAACAGGGCATCGCGGCAACGAGCAGATTCGACAAAGCTGGTGAAGCAAGGGCA 2338
Db 4163 TACCAGCACACAAGGNAATGGAGGAATGAACAAGTAGATAAATAGTCAGTCTCTGGAA 4222
QY 2339 TCCGCAAGGTGCTGTTCTCGGACGGCATCGAT 2370
Db 4223 TCAGGAAAGTACTATTTTGAATGGAATAGAT 4254

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:45:48 ; Search time 107.929 Seconds
(without alignments)
10843.369 Million cell updates/sec

Title: US-09-610-313-30
Perfect score: 2469
Sequence: 1 gtcgacgcacacattggccga.....ggcgtagcaccggtagaattc 2469

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCT_US_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US50_NEW_PUB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US50_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1659.4	67.2	4307	10	US-09-999-183-2
2	1592.8	64.5	8366	10	US-09-872-733-6
3	1557.4	63.1	4338	10	US-09-872-733-1
4	1190.6	48.2	2467	10	US-09-872-733-3
5	1132	45.8	2601	10	US-09-735-487-7
6	1132	45.8	2601	10	US-09-735-487-9
7	1132	45.8	2601	10	US-09-735-487-11
8	1132	45.8	2601	10	US-09-735-487-13
9	1132	45.8	2601	10	US-09-999-183-1
10	1132	45.8	9719	10	US-09-737-190A-1
11	1079.6	43.7	8933	10	US-09-943-286-4
12	1079.6	43.7	8933	10	US-09-943-286-5
13	1079.6	43.7	8933	10	US-09-943-286-9
14	1073.2	43.5	9544	10	US-09-798-675-4
15	1073.2	43.5	9918	10	US-09-798-675-5
16	868.8	35.2	2348	9	US-10-097-997-1
17	856.2	34.7	9793	9	US-09-886-156-56
18	856.2	34.7	9793	9	US-09-886-150-56
19	856.2	34.7	9793	9	US-09-886-149-56

20	856.2	34.7	9793	9	US-09-886-159-56	Sequence 56, Appl
21	750.2	30.4	2507	10	US-09-872-733-2	Sequence 2, Appl
22	698	28.3	12379	10	US-09-991-258-14	Sequence 14, Appl
23	694.8	28.1	1323	10	US-09-991-258-15	Sequence 15, Appl
24	475.8	19.3	948	10	US-09-756-551A-5	Sequence 5, Appl
25	243.8	9.9	1800	9	US-10-003-035-58	Sequence 58, Appl
26	239.2	9.7	4516	10	US-09-988-355-24	Sequence 24, Appl
27	216.4	8.8	12523	10	US-09-991-258-1	Sequence 1, Appl
28	215.6	8.7	1476	10	US-09-991-258-4	Sequence 4, Appl
29	188.6	7.6	9464	10	US-09-738-847-1	Sequence 1, Appl
30	179.4	7.3	13049	9	US-09-996-073-1	Sequence 1, Appl
31	179.4	7.3	13049	10	US-09-231-235-1	Sequence 1, Appl
32	179.4	7.3	13049	10	US-09-797-518A-1	Sequence 1, Appl
33	172.6	7.0	507	9	US-10-003-035-56	Sequence 56, Appl
34	163.2	6.6	1482	10	US-09-818-443-4	Sequence 4, Appl
35	163.2	6.6	1532	10	US-09-818-443-1	Sequence 1, Appl
36	157.8	6.4	1503	10	US-09-968-355-25	Sequence 25, Appl
37	156.6	6.3	1752	10	US-09-968-355-16	Sequence 16, Appl
38	156.6	6.3	4594	10	US-09-968-355-15	Sequence 15, Appl
39	152.8	6.2	9751	10	US-09-946-239-7	Sequence 7, Appl
40	133.6	5.4	1496	9	US-10-003-035-17	Sequence 17, Appl
41	116.2	4.7	2523	10	US-09-476-242-15	Sequence 15, Appl
42	115.2	4.7	2517	10	US-09-476-242-17	Sequence 17, Appl
43	112	4.5	2517	10	US-09-476-242-16	Sequence 16, Appl
44	111.8	4.5	2310	10	US-09-476-242-21	Sequence 21, Appl
45	111.8	4.5	2535	10	US-09-476-242-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-999-183-2
; Sequence 2, Application US/09999183
; Patent No. US20020147169A1
; GENERAL INFORMATION:
; APPLICANT: MITROPHANOUS, et al
; TITLE OF INVENTION: In Vivo Selection Method
; FILE REFERENCE: 674523-2009
; CURRENT APPLICATION NUMBER: US/09/999,183
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: PCT/GB00/02136
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 9912965.2
; PRIOR FILING DATE: 1999-06-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: SeqWin99
; SEQ ID NO 2
; LENGTH: 4307
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Codon optimised gagpol sequence
US-09-999-183-2

Query Match	67.2%	Score	1659.4	DB	10	Length	4307
Best Local Similarity	82.4%	Pred. No.	2.4e-266				
Matches	1956	Conservative	0	Mismatches	401	Indels	16
Gaps	4						
QY	12	CATGCCCGAGCCCATGAGCCAGG	---	CCACGAGCGCCCAACATCCTGTGATGCAGCGCAGCAA	68		
Db	1086	CCTGCTGAGGCCATGAGCCAGGTACCACTCCGCTACCATCATGATGCAGCGGGCAA	1145				
QY	69	CTTCAAGGGCCCAAGCCGCATCATCAAGTGTTCAACTGCGGCAAGGAGGCCACATCGC	128				
Db	1146	CTTTTCGGAACCAAGCAAGATCGTCAAGTGTTCAACTGTCGCAAGGAAGGACACACAGC	1205				
QY	129	CCGCAACTGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTCCGCGCAGGAGGGCCACCA	188				
Db	1206	CCGCAACTGCGAGGGCCCGCTAGGAAAAGGGCTGCTGGAATGCGGCAAGGAGGACCA	1265				
QY	189	GATGAAGGACTGCAACCGAGCGCCAGGCGCAACTTCTTCGCGAGGACCTGGCTTCCCCCA	248				

RESULT 2

US-09-872-733-6
; Sequence 6, Application US/098727233
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-42870S1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of the construct pCMVgagpolBKNK containing a CMV
; OTHER INFORMATION: promoter, a HIV gag/pol gene and a kanamycin
; OTHER INFORMATION: resistance gene
US-09-872-733-6

Query Match 64.5%; Score 1592.8; DB:10; Length 8366;
Best Local Similarity 80.7%; Pred. No. 2.5e-255;
Matches 1914; Conservative 0; Mismatches 442; Indels 16; Gaps 4;

Qy 14 TGGCGAGGCGCATGAGCGAGCCACCGC---GCCAATCTCTGATGAGCGCAGCAACT 70
Db 1857 TGGCGAGGCGATGCGCAGGTGACGACTCGGCNCCATATGTCAGAGAGCGCACT 1916
Qy 71 TCAGAGGCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGAGAGGCCACATCGCCC 130
Db 1917 TCCGGAACACGCGGAAGATCTCAAGTCTTCAATTTGGCAAGAGAGGCCACCGCCA 1976
Qy 131 GCAACTGCGCGCCCCCGCGAAGAGGCTCTGGAAGTGGCGCAAGAGAGGCCACCGAGA 190
Db 1977 GGAAGTGGCGGGCCCCCGGAAGAGGCTGTGGAAATGTGGAAAGAGAGACACAAA 2036
Qy 191 TGAAGGATGCAACGAGCGAGCCAGCGCAACTTCTTCCGCGAGGACCTGGCCCTTCCCGCAGG 250
Db 2037 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTITAGGAGAGATGCGCCTTCCCTACAG 2095
Qy 251 GCAAGGCGCGAGTTCGCCAGCGAGAGAACCGCGGCCAACAGCCCCACCGCGCGAGC 310
Db 2096 GGAAGGCCAGGGAATTTCTTTCAGAGCAGACGAGCCACAGCCCAACAGCCCAAGAGAGC 2155
Qy 311 TGCAGGTGCGGG-----CGACAACCCCCGAGCGAGCGCGCGCGCGCGCGCGCA 364
Db 2156 TTCAGGTCTGGGTAGAGAACACACATCTCCCTCTAGAAAGAGGAGCGCGATAGACAAAGAA 2215
Qy 365 -----CCCTGAACTTCCCGCAGATCACTCTGTGGCAGCGCCCTTGGTGTAGCATCAAGG 418
Db 2216 CTGTATCTTAACTTCCCTCAGATCACTCTTGGACAGACCCCTCGTCACATAGGA 2275
Qy 419 TGGCGGCGCAGATCAAGAGGCGCTGTGGACACCGCGCGAGACACCGCTGTGTGAGG 478
Db 2276 TCGGGGGGCACTCAAGGAGCGCTGCTCGATACAGGAGCAGATGATACATTAGGA 2335
Qy 479 AGATGAGCCTGCCCGCAAGTGGAGCCCAAGATGATCGCGCGCATCGCGGCTTCATCA 538
Db 2336 AATGAGTTTGGCCAGGAAGTGGAAACCAAAATGATAGGGGGATCGGGGGCTTCATCA 2395
Qy 539 AGGTCCGCCAGTACGACAGATCTGTATCGAGATCTCGGCAAGAGGCCATCGGCACCG 598
Db 2396 AGGTGAGGCGAGTACGACAGTACTCATAGAAATCTGTGGACATTAAGCTATAGTACAG 2455

Qy 599 TGCTGATCGGGCCCCCACCCCGTGAACATCATCGGGCGCAACATGCTGACCCAGCTGGGCT 658
Db 2456 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAATCTGTTGACCCAGATCGGCT 2515
Qy 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTCCCGCTGAAGCTGAAGCCCGGCA 718
Db 2516 GCACCTTGAACCTTCCCATCAGCCCTATTGAGACGGTCCCGCTGAAGTTGAAGCCCGGCA 2575
Qy 719 TGGAGCGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCGCTGACCG 778
Db 2576 TGGAGCGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCGCTTGTAGTG 2635
Qy 779 CCATCTGGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCGCAAGCCCT 838
Db 2636 AATCTGTACAGAGATGGAGAAGGAGGCAAGATCAGCAAGATCGGCCCTGAGAACCCCT 2695
Qy 839 ACAACACCCCGTGTTCGCCATCAAGAGAAGAGCAGCACCAAGTGGCCCAAGCTGGTGG 898
Db 2696 ACAACACTCCAGTCTTCCCAATCAAGAGAAGAGAGTACCAAGTGGAGAAAGCTGGTGG 2755
Qy 899 ACTTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGAGGTGCGAGCTGGGCATCCCC 958
Db 2756 ACTTCAGAGAGTGAACAGAGAACTCAGGACTTCTGGGAAGTTTCACTGGGCATCCCC 2815
Qy 959 ACCCGCGCGCGCTGAAGAAGAGAGCGTGCACCGTGTGCACGTGGCGGACGCTACT 1018
Db 2816 ATCCGCTGGGTGAAGAAGAGAGTCACTGACAGTCTGGATGTGGTGTGCTACT 2875
Qy 1019 TCAGCGTCCCTGGACGAGGACTTCCGGAAGTACACCGCTTCCACCTCCACGATCA 1078
Db 2876 TCTCGTCTCCCTTGGACGAGGACTTCCAGGAAGTACACTGCCCTTCCAGATACCTAGCATCA 2935
Qy 1079 ACAACGAGACCCCGCGCATCCGCTACCACTGACAGTACAAAGTCTGCCCGCGAGGCTGGAAGGCA 1138
Db 2936 ACAACGAGACCCCGCGCATCCGCTACCACTGACAGTACAAAGTCTGCCCGCGAGGATGGAAGGAT 2995
Qy 1139 GCGGAGCATCTTCCAGAGCAGCATGACCAAGATCTTGAGGCGCTTCCGCGCGCGCAAC 1198
Db 2996 CACGAGCATCTTCCAAAGCAGCATGACCAAGATCTTGAGGCGCTTCCGCAAGCAAGACC 3055
Qy 1199 CCGAGATCGTATCTACCATGATGAGAGCTTACGTGAGGAGCGAGCTGGAGATGG 1258
Db 3056 CAGACATCGTATCTATCAGTACATGAGGAGCTTACGTAGGAGTACCTTACCTAGGAGTGGAGATCG 3115
Qy 1259 GCGAGCAGCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGCTTCCACCA 1318
Db 3116 GCGAGCAGGACCAAGATCGAGGAGCTGAGACAGTCTGTTGAGGTGGGAGTGAACCA 3175
Qy 1319 CCGCGACAAGAACACCAAGAGGAGCGCCCTTCTGTGGATGGGCTACGAGCTGCAC 1378
Db 3176 CACGAGACAAGAACACCAAGAGGAACTTCCCTTCTGTGGATGGGCTACGAACTGCATC 3235
Qy 1379 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCGCGAGAGAGAGCTGGACCGTGAAG 1438
Db 3236 CTGACAAGTGGACAGTGCAGCCCATCTGCTGCTTGAAGAGACAGCTGGACTGTGAAGC 3295
Qy 1439 ACATCCACAAGCTGTGGGCAAGTGAACCTGGCGCAGCAGATCTACCCCGCATCAAG 1498
Db 3296 ACATACAGAGCTGTGGGCAAGTTGAATGGGCAAGCCAGATCTACCCAGCATCAAG 3355
Qy 1499 TCGCGCAGCTGTCAAGCTGTGCGCGCGCGCAAGCGCTGACCCGACATCTGTCGCCCTGA 1558
Db 3356 TTAGCAGCTGTGCAAGCTGTTCGAGGAACCAAGGCACTGACCAAGATGATCCACATGA 3415
Qy 1559 CCGAGGAGCGCGAGTGGAGCTGGCCGAGAACCGGCGAGATCTGCGGAGAGCGCGCTGCACG 1618
Db 3416 CAGAGGAGCAGAGTAGAAGTGGCAGAGAACCGGAGATCTCTGAAGGAGCGCAGTACATG 3475
Qy 1619 GCGTGTACTACGACCCAGAGAGGAGCTGGTGGCGAGATCCCGGAGATCCAGAGAGCGCGCAGCACC 1678
Db 3476 GAGTGTACTACGACCCAGAGAGGAGCTGTGTCGAGAGATCCAGAGATCCAGAGCGGGGCAAGGCC 3535
Qy 1679 AGTGGACCTACAGATCTTACCGAGGAGCGCTTCAAGAACCTGAAGACCGCGCAAGTACGCCA 1738

Db 3536 AATGGACCTACAAATCTACAGAGCCCTTAAAGACCTGAAGACAGGCAAGTACGCAA 3595
Qy 1739 AGATGGCACCGCCACACACAGAGCTGAAGCAGCTGACCGAGGCGGTGACGAAGATCG 1798
Db 3596 GGATGAGGGGTGCCACACACAGATGTGAAGCAGCTGACAGAGGCGAGTGCAGAAGATCA 3655
Qy 1799 CCATGGAGACATCGTGATCTGGGGCAAGACCCCAAGTTCCCGCTGCCCATCCAGAGG 1858
Db 3656 CCACAGAGACATCGTGATCTGGGGCAAGACTCCCAAGTTTCAAGCTGCCCATACAGAGG 3715
Qy 1859 AGACCTGGGAGACCTGGTGACCGACTACTGGCAGGCGACCTGGATCCCGAGTGGGAGT 1918
Db 3716 AGACCTGGGAGACATGGTGACCGCTACTGGCAAGCCACTGGATCCCTGAGTGGGAGT 3775
Qy 1919 TCGTGAACACCCCGCCCTGGTGAAGCTGTGGTACCAAGCTGGAGAGGAGGCCCATCATCG 1978
Db 3776 TCGTGAACACCCCTCCCTTGGTGAAGCTGTGGTATCAGCTGGAGAGGAACCCATCGTG 3835
Qy 1979 GCGCGGAGACCTTACGTGGAGCGGCGCCCAACCGCAGACCAAGATCGGAAGGCGG 2038
Db 3836 GAGCAGAGACCTTCTACGTGGATGGGCGAGCCCAACAGGGAGACCAAGCTGGCAAGCGAG 3895
Qy 2039 GCTACGTGACCGACCGGGCGGCGAGAGATCGTGAAGCTGACCGTGAAGACCAACACAGA 2098
Db 3896 GCTACGTGACCAACCGAGGACGACAGAAAGTGGTGAAGCTGACCTGACACCAACACAGA 3955
Qy 2099 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACACGCGCAGGAGTGAACATCG 2158
Db 3956 AGACTGAGCTGCAAGCCATACCTAGCTGTCAAGACACGCGACTGGAAGTGAACATCG 4015
Qy 2159 TGACCGGACGACCTAGCGCCCTGGGCATCATCCAGGCGCCGACGAGGCGAGAGCG 2218
Db 4016 TGACGAGACTCACAGTACGACTGGGCATCATCCAGGACACACAGCAATCCGAGTCA 4075
Qy 2219 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGCTGACTGAGCTGGG 2278
Db 4076 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGCTGACTGAGCTGGG 4135
Qy 2279 TGCCCGCCACAGGGGATCGCGGCAACGACGACGATCGACAGCTGTTGAGCAAGGCA 2338
Db 4136 TACAGCACACAAAGGAATTTGGAGGAATGAACAGTAGATAAATAGTCAGTCTGGGA 4195
Qy 2339 TCCGCAAGGTGCTTCTCTGACGGCATCGAT 2370
Db 4196 TCCGGAAGGTGCTTCTCTGACGGGATCGAT 4227

RESULT 3

US-09-872-733-1
; Sequence 1, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4338
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Gag/Pol gene
US-09-872-733-1

Query Match 63.1%; Score 1557.4; DB 10; Length 4338;
Best Local Similarity 80.5%; Pred. No. 1.7e-249;
Matches 1914; Conservative 0; Mismatches 441; Indels 22; Gaps 7;
Qy 14 TGCCCGAGGCGCATGAGCAGGCGCCACCAGC---GCCAACATCTCTGATCGAGCGCAACT 70
Db 1085 TGCCCGAGGCGCATGAGCAGGCTGACGAACCTCGCGACCAATAATGATCGAGAGGCAACT 1144
Qy 71 TCRAAGGGCCCCAAGCGCATCATCAAGTGTCTTCAACTCGGCAAGGAGGCGGCACATCGCCC 130
Db 1145 TCAGGAACCAAGCGAAGATCGTCAAGTGTCTTCAATTTGGCAAGAGAGGCGCACACGCCA 1204
Qy 131 GCAACTCCCGCGCCCGCCCAAGAAAGGCTGTGGAAAGTGGCGCAAGAGGCGGCACCAGA 190
Db 1205 GGAATCTCCCGGGCCCCCGGAGAGAGGCTGTGGAAATGTGGAAGAGGAGGACACCAA 1264
Qy 191 TGAAGGACTCACCGAGCGCGCCAGCCAACTTCTTCCCGAGGAGCTGGCTTCCCCCAGG 250
Db 1265 TGAAGATTTGACTGAGAGACAGGCTAA-TTTTTTAGGGAAGATCTGGCTTCTCTACAAG 1323
Qy 251 GCAAGGCGCGGAGTTCCCGCAGCAGCAGACACCGCCCAACAGCGCCACCGCCGCGAGC 310
Db 1324 GGAAGGCGCAGGGAATTTTCTCAGAGCAGACACAGAGCCACAGCCCGCCACGAAGAGAGC 1383
Qy 311 TGACAGTGCAGCGG-----CGACAAACCCCGCAGCAGGCGCGCGAGCGCGCGAGCGCGCA 364
Db 1384 TTCAAGTCTGGGTAGAGACAACAACCTCCCGCTCAGAAGCAGGAGCGGATAGACAAGAA 1443
Qy 365 -----CCCTGAACTTCCCGCAGATCACTCTGTGGCAGCGCCCGCTGTGAGCATCAAG 418
Db 1444 CTGTATCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCTGTACAGATAAG 1503
Qy 419 TGCGCGCCAGATCAAGGAGCGCTCTGTGACACCGCGCGCGCAGCAGACCGTGTGGAGG 478
Db 1504 TCGGGGGCGCACTCAAGGAAGCGCTCTGATACAGGAGCAGATGATACAGTATTAGAG 1563
Qy 479 AGATGAGCCTCCCGCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
Db 1564 AATGAGTTTGGCAGGAAGATGGAACCAAAATGATAGGGGGATCGGGGCTTCATCA 1623
Qy 539 AGTGCGCCAGTAGCAGCAGATCTCTGATCGAGATCTCGGCAAGAGGCCATCGCACCG 598
Db 1624 AGGTGAGCGAGTAGCAGCAGATCTCATAGAAATCTGTGGACATATAAGCTATAGTACAG 1683
Qy 599 TGCTGATCGCGCCCGCACC---CCGTTGAACATCATCGCGCGCAACATGCTGACCCAGCTG 654
Db 1684 TATTAGTAGGACCTTACCTTACACCTGTCAACATAATTGGAAGAAATCTGTTGACCCAGATC 1743
Qy 655 GGCTGCAACCTTGAACCTTCCCGCATCAGCGCCCATCGAGACCGTGCCTGTAAGTGAAGCCC 714
Db 1744 GGCTGCACTTGAACCTTCCCGCATCAGCCCTATTGAGAGCGGTGCCCGTGAAGTGAAGCG 1803
Qy 715 GGCATGAGCGCCCGCAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTG 774
Db 1804 GGCATGAGCGCCCGCAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTTA 1863
Qy 775 ACCGCCATCTCGAGGAGATGGAGAAGGAGGCGCAAGATCACCAAGATCGCGCCCGCAGAAC 834
Db 1864 GTCGAATCTGTACAGAGATGGAGAAGGAGGAGATCAGCAAGATCGGCGCTTGAGAAC 1923
Qy 835 CCCTTACAACACCCCGCTGTTCGCCATCAAGAAGAGGAGACACCAAGTGGCGCAAGCTG 894
Db 1924 CCCTTACAACACTTCCAGTCTTCGCAATCAAGAAGAGGAGACAGTACCAAGTGGGAAAGCTG 1983
Qy 895 GTGACCTTCCCGAGCTGAACAAAGCGCACCGACCTTCTGGGAGGTGCGATCGGGCATC 954
Db 1984 GTGACCTTCAAGAGCTGAACAAAGAGAACTCAGGACTTCTGGGAGGTTCAGCTGGGCACT 2043
Qy 955 CCCACCGCGCGCGCTGAAGAAGAGAGCGCTGACCGCTGGAGCTGGGCGACGCC 1014
Db 2044 CCACATCCCGTGGTGAAGAAGAGAAAGTCAAGTACAGTGTGGTGTGATGCC 2103

QY	1015	TACTTTCAGCTGCCCTGGACGAGGACTTCCGCAAGTACACGGCTTACACATCCCCAGC	1074
DB	2104	TACTTCTCCGTTCCCTTGGACGAGGACTTTCAGGAAGTACACTGCTTCACGATACCTAGC	2163
QY	1075	ATCAACAACGAGACCCCGCATCCGCTACCACTACACGTGCTTCCCCAGGGCTGCAAG	1134
DB	2164	ATCAACAACGAGACACAGGATCTCCCTACCAAGTACACGTGCTGCCACAGGATGGAAG	2223
QY	1135	GGACGCCCAAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCGC	1194
DB	2224	GGATCACCCACCATCTTTCAAAGCAGCATGACCAAGATCTCTGGAGCCCTTCCCGCAAGCAA	2283
QY	1195	AACCCCGAGATCGTGTACTACCAAGTACATGGAGACCTTGTACGTGGGACGACCTGGAG	1254
DB	2284	AACCCAGACATCGTGTACTATCAGTACATGAGCAGACCTCTACGTAGGAAGTGTAGCTGGAG	2343
QY	1255	ATC-GGGCAGCACCGGCCAAGATCGAGAGCTGCGCAAGCACCTGCTGCCTGGGGCTT	1313
DB	2344	ATCGGGCAGCAGAGCAACCATGATCGAGAGCTGAGACAGCATCTGTGTAGGTGGGACT	2403
QY	1314	CACACCCCGCACAAGAAGCACCAAGAAGAGCCGCCCTTCTGTGGATGGGCTACGAGCT	1373
DB	2404	GACCACACCAAGAAGCACCAAGAAGAACCTTCCCTTCTGTGTGGATGGGCTACGACT	2463
QY	1374	GCACCCCGCAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAAAGAGAGCTGGACGT	1433
DB	2464	GCATCTTGACAAGTGGACAGTGCAGCCATCTCTGTGCTGAGAAAGACAGCTGGACTGT	2523
QY	1434	GAACGACATCCAGAAGCTGGTGGCAAGCTGAACCTGGCCAGCCAGACATCTACCCCGGCAT	1493
DB	2524	GAACGACATCAGAGAGCTCGTGGCAAGTGAACCTGGCAAGCCAGAGATCTACCCAGGCAT	2583
QY	1494	CAAGGTGCGCAGCTGTGCAAGCTGTGCGCGCGCCCAAGCCCTGACCGACATCTGTGC	1553
DB	2584	CARAAGTTAGCAGCTCTGCAAGCTGCTTCGAGAACCAAGGACCTGACAGATGATCCC	2643
QY	1554	CTGTACCGAGAGCGCAGCTGGAGCTGGCGAGAAACCGAGAGATCTGTGCGGAGCCCGT	1613
DB	2644	ACTGACAGGAAGACAGAGCTAGAATGSCAGAGAACCGAGAGATCTCTGAAGGAGCCAGT	2703
QY	1614	GCAGGGCTGTACTACGACCCAGCAAGAGCTGTGTGGCCGAGATCCAGAGCAGGGCCA	1673
DB	2704	ACATGGAGTGTACTGACCCAGCAAGAGCTGATGTCAGAGATCCAGAGCAGGGGCA	2763
QY	1674	CGACAGTGGACCTTACAGATCTACAGAGCGCTTCAAGAACCTGAAGACCCGCGCAAGTA	1733
DB	2764	AGGCCAATGGACCTACCAATCTACAGGAGCCCTTCAAGAACCTGAAGACAGGCAAGTA	2823
QY	1734	CGCCAGATGCGCACCGCCCAACCAACAGAGTGAAGCAGCTGACCGAGGCCCTGTCAGAA	1793
DB	2824	CGAAGGATGAGGGTGCACACCAACCAAGATGTGAAGCAGCTGACAGAGCAGTGCAGAA	2883
QY	1794	GATCGCATGGAGAGCATGTTATCTGGGGCAAGACCCCAAGTTTCGCTTCCGCTATCCA	1853
DB	2884	GATCACCAAGAGAGCATGTTATCTGGGGCAAGACTTCCCAAGTTCAAGTGTGCCCATACA	2943
QY	1854	GAAGGAGACTGGGAGACCTTGTGTGACCGACTACTTGGCAGGCCACCTTGTATCCCGAGTG	1913
DB	2944	GAAGGAGACATGGAGACATGTTGGACCCAGTACTGGCAAGCCACCTTGGATCCCTGAGTG	3003
QY	1914	GGAGTTGTGAACACCCCCCCTTGGTGAAGCTGTGTGTGTTACAGCTGGAGAGAGGCCAT	1973
DB	3004	GGAGTTGTGAACACCCCCCTTGGTGAAGCTGTGTGTGTTATCAGCTGGAGAGGAACCCAT	3063
QY	1974	CATCGGGCGGAGACCTTCTAGTGGAGCGGCGCCCAACCGCAGACCAAGATCTCGGCAA	2033
DB	3064	CGTGGGAGAGAGACCTTCTAGTGGATGGGGCAGCCCAACAGGAGACCAAGTGTGGCAA	3123
QY	2034	GGCGGCTACGTGACCGACCGGGCGCGCAAGATCTGTAGCTGTGACCGACCAACCAA	2093
DB	3124	GGCAGGCTACGTGACCAACCGAGAGCAGACAGAAAGTGGTGAACCTGACTGACACCAACAA	3183
QY	2094	CCAGAAGCCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCAAGAGGTGA	2153

Db 3184 CCAGAAGACTGAGCTGCAAGCCATCTACCTAGCTCTCTCAAGACAGCGGACTTGAAGTGAA 3243
QY 2154 CATCGTGACCGACAGCCAGTAGCCCTGGGCATCATCCAGGCCCGCCGCAAGAGCGGA 2213
Db 3244 CATCGTGACAGACTCAACAGTAGC-CATGGGGCATCATCCAAGCACACCGACCACTCCGA 3302
QY 2214 GAGCGAGCTGGTGAACAGATCATCGACAGCTGATCAAGAAAGGAGAAGTGTACCTGAG 2273
Db 3303 CTCAGAGCTGGTGAACAGATCATCGACAGCTGATCAAGAAAGGAGAACTGTACCTGCC 3362
QY 2274 CTGGGTGCCGCCCCACAGGCGCATCGCGGCGACAGCAGCATCGACAAGCTGGTGAGCAA 2333
Db 3363 ATGGGTACCGACACAAAGGAATTGGAGGAATGAACAAGTAGATAAATTAGTCAGTGC 3422
QY 2334 GGGCATCGCAAGTGCTGTTCCTGGAGCGGCATCGAT 2370
Db 3423 TGGGATCGGAAGTGCTGTTCCTGGAGCGGCATCGAT 3459

RESULT 4
US-09-872-733-3
; Sequence 3, Application US/09872733
; Patent No. US2001003665A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, HIV GAG AND
; TITLE OF INVENTION: HIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, HIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; US-09-872-733-3

Query Match	48.2%	Score 1190.6	DB 10	Length 2467
Best Local Similarity	84.6%	Pred. No. 7e-189		
Matches 1337	Conservative 0	Mismatches 244	Indels 0	Gaps 0
Qy 790	GAGATGGAGAGGAGGGCGAAGATCACCAAGATCGGCCCGGAGAACCCCTACAAACACCCCC	849		
Db 7	GAGATGGAGAGGAGGGAAGATCAGCAAGATCGGGCCTGAGAACCCCTACAACTACCA	66		
Qy 850	GTGTTTCGCCATCAAGAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTTCGCGGAG	909		
Db 67	GTCTTCGCAATCAGAAGAAGACAGTACCAAGTGGAGAAAGCTGGTGGACTTCAGAGAG	126		
Qy 910	CTGAAACAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCGCCACCCCGCGCGC	969		
Db 127	CTGAACAAGAGAACTCAGGAGCTTCTGGGAAGTTCAGCTGGGCATCCACATCCCGCTGGG	186		
Qy 970	CTGAAGAAGAGAAGAGCGTGACCGTGTGCAGCTGGGCGAGCGCCTACTTTCAGCGTGGCC	1029		
Db 187	TTGAAGAAGAAGAAGTCAGTGACAGTGTGTGATGGGTGATGTGCTACTTTCGGTTCGCC	246		
Qy 1030	CTGGACGAGGACTTCCCGCAAGTACACCGCCCTTACCATTCCCGAGCATCAACAACGAGACC	1089		
Db 247	TTGGAGGAGGACTTCAGGAAGTACACTGCCTTTCAGTATACCTAGCATCAACAACGAGACA	306		
Qy 1090	CCCGGCATCCCGTACCAGTACAAAGTGTGTCGCCCGAGGGCTGGAGGGCGAGCCCGAGCATC	1149		
Db 307	CCAGGCATCCGCTACCAAGTACAAAGTGTGTCACAGGGATGAAGGGATCACCAGGCATC	366		

QY 1150 TTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGGCCCGCAACCCCGAGATCGTG 1209
Db 367 TTTCRAAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGAAGCAAAACCCAGCATCGTG 426
QY 1210 ATCTACAGTACATGACGACACCTGTACGTGGCAGGACCTGGAGATCGGCGACGACCCG 1269
Db 427 ATCTATCAGTACATGACGACACCTGTACGTAGGAAGTGACCTGGAGATCGGCGACGACAG 486
QY 1270 GCAAGATCGAGAGCTGGCGCAAGACCTGTCTGGCTGGGGCTTACACACCCCGACAAAG 1329
Db 487 ACCAAGATCGAGAGCTGGACACATCTCTGTTGAGGTGGGGAGTACACACACCGACAAAG 546
QY 1330 AAGCAGCAAGAGGAGCCCTCTCTGTGTGATGGGTACGAGCTGCACCCCGACAAAGTG 1389
Db 547 AAGCAGCAAGAGGAGCCCTCTCTGTGTGATGGGTACGAGCTGCACCTGTGACAAAGTG 606
QY 1390 ACCGTGCAGCCCATCGAGCTGCCGAGAGAGGAGTGGACCTGGAACGACATCCAGAAAG 1449
Db 607 ACAGTGCAGCCCATCGTGTCTGCTGAGAGGACAGCTGGACTGTGAAGGACATACAGAAG 666
QY 1450 CTGGTGGCAAGCTGAACCTGGCCAGCAGCATCTACCCGGGATCAAGGTGCCCGACGCTG 1509
Db 667 CTGGTGGCAAGCTGAACCTGGCCAGCAGCATCTACCCGGGATCAAGGTGGCGAGCTG 726
QY 1510 TGCAAGCTGTCTGGCGGCGCAAGGCCCTGACCGACATCTGCTGCCCTGACCGAGGAGGCC 1569
Db 727 TGAAGCTGTCTGAGGAACCAAGGCACTGACAGAGTGTATCCACTGACAGAGGAGCA 786
QY 1570 GAGCTGAGCTGTCCGCGAGAACCCGCGAGATCTCTGCGAGGCCCTGACGCGCTGTACTAC 1629
Db 787 GAGCTGAGCTGTCCGCGAGAACCCGCGAGATCTCTGAGGAGCCAGTACATGAGTGTACTAC 846
QY 1630 GACCCAGCAAGGACCTGTGGCGGAGATCCAGAGCAGGCGCCACGACCTGAGCTTAC 1689
Db 847 GACCCAGCAAGGACCTGTGCGAGAGATCCAGAGCAGGCGCCAGGCGCAATGGACCTAC 906
QY 1690 CAGATCTACGAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCAAGATGCGCAC 1749
Db 907 CAATCTACGAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCAAGATGAGGGGT 966
QY 1750 GCCACACCAAGCAGCTGGAAGCAGCTGACGAGCGCGTGCAGAAATGCGCATGGAGAGC 1809
Db 967 GCCACACCAAGCAGCTGGAAGCAGCTGACGAGCGAGTGCAGAAATGCGCATGGAGAGC 1026
QY 1810 ATCGTGTCTGGGCGAGACCCCGCAAGTTCGGCTGCCCATCCAGAGGAGACCTGGAG 1869
Db 1027 ATCGTGTCTGGGCGAGACCCCGCAAGTTCGGCTGCCCATCCAGAGGAGACATGGGAG 1086
QY 1870 ACCTGTGGAGCCGACTACTGGCAGGCGCACCTGGATCCCGGAGTGGGATTCGTGAACACC 1929
Db 1087 ACATGTGGAGCCGACTACTGGCAGGCGCACCTGGATCCCGGAGTGGGATTCGTGAACACC 1146
QY 1930 CCCCCCTGTGTGAGCTGGTACAGCTGGAGAGGAGCCCATCATCTGGCGCGGAGACC 1989
Db 1147 CTTCTCTGTGTGAGCTGGTATCAGCTGGAGAGGAGCCCATCATCTGGCGCGGAGACC 1206
QY 1990 TTCTACTGTGACGCGCGCCGACCGGAGACCAAGATCGGAGGCGGCTACGTGACC 2049
Db 1207 TTCTACTGTGAGGCGGCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1266
QY 2050 GAGCGGGCGCGGAGAGATCTGTGAGCTGTGACCGGAGACCAACCAAGAGACCGAGCTG 2109
Db 1267 AAGCGGAGGAGAGAGAGTGTGTGACCTGACCTGACACCAACCAAGAGAGTGTGAGCTG 1326
QY 2110 CAGGCGATTCAGCTGGCGCTTGCAGGAGCAGCGGAGGAGTGAACATCTGTGACCGAGC 2169
Db 1327 CAGGCGATTCAGCTGTGCAAGAGAGGAGGAGTGAACATCTGTGACAGAGTCA 1386
QY 2170 CAGTACGCGCTGGGCGATCATCCAGGCGCCAGCGACAGAGGAGGAGGAGGAGGAGGAG 2229
Db 1387 CAGTACGCGCTGGGCGATCATCCAGGCGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1446

QY 2230 CAGATCATCGAGCAGCTGATCAAGAAGGAGAGTGTACTGAGCTGGGTGCCGCCAC 2289
Db 1447 CAGATCATCGAGCAGCTGATCAAGAAGGAGAGTGTACTGAGCTGGGTACCAGCAC 1506
QY 2290 AAGGGATCGGCGGCAACGAGCAGATCGACAAGCTGTGAGCAAGGCGATCCGCAAGGTG 2349
Db 1507 AAAGGAATCGGAGAAATGAACAAGTAGATAAATAGTCAAGTCTGGGATCCGGAAGGTG 1566
QY 2350 CTGTTCTCGACGGCATCGAT 2370
Db 1567 CTGTTCTCGACGGGATCGAT 1587
RESULT 5
US-09-735-487-7
; Sequence 7, Application US/09735487
; Patent No. US20020042679A1
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/735,487
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/117,217
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(492)
; OTHER INFORMATION: gag Polyprotein
US-09-735-487-7

Query Match 45.8%; Score 1132; DB 10; Length 2601;
Best Local similarity 68.5%; Pred. No. 3.4e-179;
Matches 1626; Conservative 0; Mismatches 730; Indels 16; Gaps 4;

QY 14 TGCGCGAGGCGCATCAGCGCATCATCAAGTGTCTCAACTCGCGCAAGGAGGGCCACATCGCC 70
Db 77 TGGCTGAAGCAATGAGCAAGTACAAATTCAGCTACCAATATGATCGAGAGGCAATT 136
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTCGCGCAAGGAGGGCCACATCGCC 130
Db 137 TTAGGAACCAAGAAAGATGTTAAGTGTTCATATTCTGGCAAGAAAGGGCCACAGCCA 196
QY 131 GCAACTCGCGGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGGGCAAGAGGGCCACAGAG 190
Db 197 GAAATTCAGGGCGCCCTAGGAAAAGGGCTGTTGGAATTTGGAAGAAAGAGGACCAAAA 256
QY 191 TGAGAGGACTCAGCGAGCGCGCAGCCCAACTTCTTCGCGAGGAGCTGGCTTCCCGCAGG 250
Db 257 TGAAGAGATTGTACTGAGAGACAGGCTAA-TTTTGTAGGAGAGATCTGGCTTCTTCAAG 315
QY 251 GCAAGGCGCGGAGTTCCTCCAGCGCAGCAGAACCCGCGCAACAGCCCAACAGCGCGAGC 310
Db 316 GGAAGGCGAGGAATTTTTCAGAGCAGACACAGAGCCACAGCCCAACAGAGAGAGC 375
QY 311 TGAGAGTGGCGG-----CGACAAACCCCGCAGCGAGCGCGCGCGAGCGCGAGGCA 364
Db 376 TTAGAGTCTGGGGTAGAGACAACAACCTCCCTTCAGAGCAGGAGGAGCGGATAGACAAGAA 435
QY 365 -----CCCTGAACCTCCCGCAGATCACTCTGTGGCAGCGCCCTGTGTGAGCATCAAG 418
Db 436 CTGTATCTTTAACTTCCCTCAGGTCACTCTTTTGGCAACGACCCCTCTGTCAATNAAGA 495

QY 419 TGGCGGCGCAGATCAAGAGGCGCCCTGTGGACACCGGCGCGACGACACCCGCTGCTGGAGG 478
Db 496 TAGGGGGGCACTAAAGAGAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 555
QY 479 AGATCAGCTGCGCGCAGTGGAGCCCAAGATGATCGGCGGCGATCGGCGGCTTCATCA 538
Db 556 AAATGAGTTTGGCAGGAGATGGAACCAAAATGATAGGGGAATTTGAGGTTTATCA 615
QY 539 AGGTGCGCAGTACACCAAGTCCCTGATCGAGATCTGGGGCAAGAGGCCATCGGCACCG 598
Db 616 AAGTAAGACAGTATGATCAGATACATAGAAATCTGTGGACATAAGACTATAGGTACAG 675
QY 599 TGCTGATCGGCGCCACCGCGTGAACATCATCGGCGGCAACATGCTGACCCAGCTGGCT 658
Db 676 TATTAGTGGACCTACACTGTCAACATAAATTGGAAGAAATCTGTGACTCAGATTGGTT 735
QY 659 GCACCTGAACCTCCCATCATCGCCCATCGAGACCGTGCCTGTGAAGCTGAAGCCCGGCA 718
Db 736 GCACCTTAAATTTCCCATTAGCCCTATTGAGACTGTACCAAGTAAATTAAGCCAGGAA 795
QY 719 TGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 796 TGGATGGCCCAAGGTTAAACAATGGCCATTCACAGAGAAATAAAGCATTAGTAG 855
QY 779 CCATCTGCGAGAGATGAGAGAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT 838
Db 856 AAATTTGTACAGAGATGGAAGAGGAGGCAAAATTTCAAAATTTGGGCTGAAATCCAT 915
QY 839 ACAACACCCCGTGTCCGCATCAGAGAGAGGAGCAGACACCAAGTGGCGCAGCTGGTGG 898
Db 916 ACAATCTCCAGTATTGCCATAAGAAAAAGACAGTACTAAATGGAGAAAAATTAGTAG 975
QY 899 ACTTCGGGAGGTGAACAGGCGACCCAGGACTTCTGGAGGTGCAGCTGGGCATCCGCC 958
Db 976 ATTTTCAGAGAACTTAATRAGAACTCAGACTCTCGGAAGTTCATTAAGGAATACCAC 1035
QY 959 ACCCGCGCGGCTGAAGAGAGAGAGCGGTGACCGCTGCTGACGCTGGCGCAGCGCTACT 1018
Db 1036 ATCCGCGAGGCTTAAAGAAAAAATCAGTAACAGTACTGGAATGGAGAAAAATTAGTAG 1095
QY 1019 TCAGGCTGCCCTGGACGAGGACTTCCGCAAGTACAGCCGCTTACCATCCCGACATCA 1078
Db 1096 TTTCAGTTCCCTTAGATCAGAGACTTCAGAGAGTATCTGCAATTTACCATACCTAGTATAA 1155
QY 1079 ACAACGAGACCCCGGCTACCGCTACAGTACAAAGCTGCTGCCCGAGGCTGGAGGGCA 1138
Db 1156 ACAATGAGACACACAGGGATTAGATATCAGTACAATGTCTTCCACAGGATGGAAAGAT 1215
QY 1139 GCGCCAGCATCTCCAGAGCAGATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACC 1198
Db 1216 CACCAAGCAATATCCAAAGTACATGACAAAAATCTTAGAGCCCTTTAGAAAAAATC 1275
QY 1199 CCGAGATCGTATCTACCAAGTACATGGAGACCTCTAGCTGGGACGACCTGGAGATCG 1258
Db 1276 CAGACATAGTTTATCTATCAATACATGGATGATTTGATGTAGGATCTGACTTAGAAATAG 1335
QY 1259 GCCAGCACCGCGCCAAAGATCAGGAGCTGCGCAAGCACTGCTGGCTGGGCTTCACCA 1318
Db 1336 GGCAGCATAGAACAAAAATAGAGGAGCTGAGACAACTCTGTGTAGGTGGGACTTACCA 1395
QY 1319 CCGCGACAAAGAACACAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1378
Db 1396 CACCAAGAAAAACATCAGAGAAACCTCCATCTCTTTGGATGGGTTATGAATCCCATC 1455
QY 1379 CCGACAAGTGACCTGACAGCCATCGAGCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGG 1438
Db 1456 CTGATAATGGACAGTACAGCCTATGAGTGTGCGCAGAAAAAGACAGCTGGACTGTCAATG 1515
QY 1439 ACATCCAGAGCTGTGTGGGCAAGCTGAATCTGGGCGAGCAGATCTACCCCGGCTCAAGG 1498
Db 1516 ACATACAGAGCTTAGTGGGGAATGANTTGGCAAGTCAGATTTACCCAGGATTAAG 1575
QY 1499 TGGCGCAGCTGTGAAGCTGTGCGCGCGGCAAGGCCCTGACCGACATCTGTCGCCCTGA 1558

Db 1576 TAAGCAATATTATTAACCTCTTAGAGGAACCAAGCACTAACAAGATTAATACCACTAA 1635
QY 1559 CCGAGGAGCCGAGCTGGAGCTGGCGGAGAACGGGAGATCTCGCGGAGCCGCTGCACG 1618
Db 1636 CAGAGAAGCAGAGTAGAACTGGCAGAAACAGAGAGATTTCTAAAGAAGAACCACTACATG 1695
QY 1619 GCGTGTACTACGACCCCAAGGACCTGCTGCCGAGATCCAGAGAGGAGGCGGCGACGAC 1678
Db 1696 GASTGTATTATGACCCATCAAGACTTAATAGCAGAAATACAGAAGCAGGGCAAGGCC 1755
QY 1679 AGTGACCTTACAGATCTTACCAGGACCTTCAAGAACCTGAAGACCCGCAAGTACGCCA 1738
Db 1756 AATGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAACACAGGAAAAATGCAA 1815
QY 1739 AGATGGCAGCCCGCCACACCAAGCAGCTGAAGCAGTGAACGAGGCGCTGCGAAGATCG 1798
Db 1816 GAATGAGGGTGGCCACACATTAATGATGTAACAAATTAACAGAGGAGTGCAAAAATAA 1875
QY 1799 CCATGGAGAGCTGCTGATCTGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAAGG 1858
Db 1876 CCACAGAAAGCATAGTATATGGGAAGAGCTCTTAATTTAACTGCCCATACAAAAG 1935
QY 1859 AGACCTGGGAGACCTGCTGGACCCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAT 1918
Db 1936 AAACATGGGAAACATGCTGGACAGAGTATTGGCAAGCCACCTGGATCTCTGAGTGGGACT 1995
QY 1919 TCGTGAACACCCCGCCCTGGTGAAGCTGTTGACAGCTGGAGAGGAGGAGGAGGAGGAGG 1978
Db 1996 TTGTTAATACCCCTCCCTTAGTGAATTTATGTTACAGTTAGAGAAAGAACCCATAGTAG 2055
QY 1979 GCGCGGAGACCTTCTACGTGGAGCGCGCCCAAGCGGAGACCAAGATCGGCAAGGCGG 2038
Db 2056 GAGCAGAAACCTTCTATGTAGATGGGAGAGCTTACAGGAGAGCTAATTAGGAAAAAGCAG 2115
QY 2039 GCTAGCTGACCGACCGCGGCGCGGAGAGATCGTGAGCTTACCAGGAGGAGGAGGAGGAGG 2098
Db 2116 GATATGTTACTAATAGAGGAAGACAAAAAATTTGTACCCCTAAGTACACAAACAAATCAGA 2175
QY 2099 AGACGAGCTGCGAGGCCATCCAGCTGGCCCTCGAGGAGGAGGAGGAGGAGGAGGAGG 2158
Db 2176 AGACTGAGTTACAGCAATTTATCTAGCTTTCAGGATTCGGGATTTAGAAAGTAAACATAG 2235
QY 2159 TGACGAGACGCGAGTACGCGCTGGGCATCATCCAGGCCAGCGCCAGAGAGAGGAGAGG 2218
Db 2236 TAACAGACTCAAAATATGATTTAGGAACTATTCAGGACCAACAGATCAAGTGAATCAG 2295
QY 2219 AGCTGGTGAACAGATCATCTGAGCAGCTGATCAAGAAGGAGAGGAGTGTACCTGAGCTGG 2278
Db 2296 AGTTAGTCAATCAAAATAATAGAGCAGTTAATAAAAAAGGAAAGGCTTATCTGGCATGG 2355
QY 2279 TGCCCGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2338
Db 2356 TACCAAGCAGCAAGGAAATTTGAGGAAATTTGAACAAAGTATAGTAAATAGTGCCTGGAA 2415
QY 2339 TCCGCAAGTGTCTTCTGGAGCGCATCGAT 2370
Db 2416 TCAGGAAGTACTATTTTTAGATGGAATAGAT 2447

RESULT 6

US-09-735-487-9
; Sequence 9, Application US/09735487
; Patent No. US20020042679A1
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/735,487


```
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type I
US-09-999-183-1

Query Match      45.8%; Score 1132; DB 10; Length 4307;
Best Local Similarity 68.5%; Pred. No. 3.5e-179;
Matches 1626; Conservative 0; Mismatches 730; Indels 16; Gaps 4;

QY 14 TGCCGAGGCGATGACGCGAGCGCCACCA--GCGCCCAACATCTTGATGCGAGCGCAACT 70
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1088 TGCTGAAGCAATGAGCCAGTAACAAATTCAGTACCATAATGATGCGAGAGGCAATT 1147

QY 71 TCAAGGCCCCCAAGCGATCATCAAGTGTCAACTGCGGCAAGGAGGCGCACATCGCCC 130
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1148 TTAGGAACCAAGAAAGATGTTAAGTGTTCATTTGTCGAAGAAAGGCGCACAGCCA 1207

QY 131 GCAACTGCCGCGCCCCCGCAAGAGGCTGCTGGAAGTCCGCAAGGAGGCGCACCA 190
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1208 GAAATTCGAGGGCCCCCTAGGAAAGAGGGCTGTGGAAATGTGGAAGAGGAGACCCAA 1267

QY 191 TGAAGACTGCACCGCGCGAGCGCCAGCAACTTCTTCGCGAGGAGCTGGCCTTCCCGCAGG 250
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1268 TGAAGATGTACTGAGAGACAGCTAA-TTTTATTAGGNAAGATCTGGCCTTCTTACAG 1326

QY 251 GCAAGGCCCGGAGTTCGCCAGCGAGCAGACGACCGCGCCCAACGCCCCACACGCGGAGC 310
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1327 GGAAGGCCAGGGAAATTTCTTTCAGAGCAGACCAAGAGCAGACGCCCCACAGAAAGAGC 1386

QY 311 TGCAGGTGGCGG- - - - -CGACAACCCCGCAGCGAGCGCGCGGAGCGCGAGGGCA 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1387 TTCAGGTCTGGGTAGAGACAACAATCCCCCTCAGAGCAGGAGCGCGATAGCAAGGAA 1446

QY 365 - - - - -CCCTGAATCTCCCCAGATCACCTGTGGCAGGCGCCCTTGGTGAGATCAAGG 418
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1447 CTGTATCTTTAACTTCCCTCAGGTCACTCTTTTGGCAACGACCCCTCGTCACATAAGA 1506

QY 419 TGGCGCGCAGATCAAGGAGGCCCTCTGTGACACCGCGCGGACGACACCGTCTGGAGG 478
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1507 TAGGGGGGCAACTTAAAGAGAGCTCTATTAGATACAGGACGAGATGATACAGATTAGAAG 1566

QY 479 AGATGAGCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGCGGCACTGGCGGCTTCATCA 538
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1567 AATGAGTTTGGCAGGAAGATGGAACCAACCAAAATGATAGGGGAATTTGAGGTTTATCA 1626

QY 539 AGTGCGCCAGTACGACCATCTCTGATCGAGATCTCGCGCAAGAGGCGCATCGGCACCG 598
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1627 AAGTAAGACAGTATGATCAGATCTCATAGAAATCTGTGGACATAAAGCTATAGGTACAG 1686

QY 599 TGCTGATCGGCCCCACCCCGTGACATCATCGGCCGCAACATGCTGACCCAGCTGGCT 658
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1687 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTGACTCAGATTGGTT 1746

QY 659 GCACCTGGAATTCGCCATCAGCCCATCGAGACCGTGCCTGGAAGCTGGAAGCCCGCA 718
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1747 GCACTTTAAATTTCCCATTTAGCCCTATTGAGACTGTACAGTAAATTAAGCCAGGAA 1806

QY 719 TGGACGCCCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCGCTGACCG 778
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1807 TGGATGCCCAAAAGTTAAACAATGCCATTGACAGAGAAAAAATAAAGCATTTAGTAG 1866

QY 779 CCATCTGCGAGAGATGGAGAGGAGGGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT 838
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1867 AAATTTGTACAGAGATGGAAAAGGAGGAAAAATTTCAAAAATTTGGCCCTGAAAATCCAT 1926

QY 839 ACAACACCCCGCTGTTCGCCATCAAGAAAGACAGCACCAAGTGGCGCAAGCTGTGG 898
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1927 ACAATCTCCAGTATTGGCCATAAGAAAAAAGACAGTACTAATGGAGAAAATTTAGTAG 1986

QY 899 ACTTCGCGAGGTGAACAGCGCACCCAGGACTTCTGGAGGTGCAAGCTGGGCATCCCCC 958
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1987 ATTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTTAGGAATACCAC 2046

QY 959 ACCCCCGCGGCTGAAGAGAGAGAGAGCGGTGACCGTGTGACGTGGCGGAGCGCTACT 1018
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```



```
QY 1499 TGGCCAGCTGTGCAAGCTGCTGGCGCGGCCAAGGCCCTGACCGACATCGTGCCCTGA 1558
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3376 TAGGCAATTATGTAAACTCTTATAGAGAACCAAGACCTAAACAGAGTAATACCACTAA 3435
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1559 CCGAGAGCGCGAGCTGGAGCTGCCCGAGAACCCGCGAGATCTTGGCGAGCCCGTGCAAG 1618
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3436 CAGAAGAGCAGAGCTAGAATCTGCAGAAAACAGAGAGTTCTAAAGAACCCAGTACATG 3495
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1619 GCGTGTACTACGACCCAGCAAGACCTGCTGGCGGAGATCCAGAACGCGGCCACGACC 1678
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3496 GAGTGTATTATGACCCATCAAGACTTAATAGCAGAAATACAGAGCAGGGCGAAGGCC 3555
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1679 AGTGGACCTACAGATCTACAGAGGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1738
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3556 AATGGACATATCAAAATTATCAGAGCCATTTAAAAATCTGAAAAACAGGAAATATGCAA 3615
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1739 AGATGCCACCGCCACACACGACGCTGAAGCAGCTGACCGAGGCGCGTGCAGAAATCG 1798
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3616 GAATGAGGGGTGCCACACTAATGATGTAAACAATTTAAACAGAGGCAAGTGCAAAAATAA 3675
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1799 CCATGGAGACATCGTATCTGGGGCAAGACCCCAAGTTCCGCCTGCCCATCCAGAAGG 1858
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3676 CCACAGAAACATAGTAATATGGGAAGACTCTTAATTTAACTGCCCATCAAAAGG 3735
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1859 AGACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1918
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3736 AAACATGGGAACATGGTGACAGAGTATTGGCAAGCCACTGGATTCTCTGAGTGGGAGT 3795
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1919 TCGTGACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGGAGAGGAGCCCATCATCG 1978
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3796 TTGTTAATACCCCTCCCTTAGTGAATATTGTTACCAAGTTAGAGAAAGCAATAGTAG 3855
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1979 GCGCCGAGACCTTCTAGCTGGAGCGCGCCCAACCCGAGACCAAGATCGCAAGCGCG 2038
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3856 GAGCAGAACTTCTATGTATAGTGGCGACCTACAGGGAGACTAATATTAGGAAAGCAG 3915
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2039 GCTACGTGACCGACCGGGCGGCAGAAATCGTGAAGCTTGACCGACACCAACCCAGA 2098
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3916 GATATGTTACTAATAGAGGAAGACAAAAGTTGTACCCCTAACTGACACAAACAATCAGA 3975
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2099 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACACGCGGAGGTGAACATCG 2158
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3976 AGATGAGTTACAGCAATTTATCTAGCTTTGAGGATTGCGGATTAGAAATTAACATAG 4035
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2159 TGACCGACACGCCAGTACGCCCTGGGATCATCCAGGCCCGCCAGCAGAGCGAGCG 2218
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4036 TAACAGACTCACATATGCATTAGGAATCATTCAGCACACCCAGATCAAGTGAATCAG 4095
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2219 AGCTGGTGAACAGATCATCGACAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 2278
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4096 AGTTAGTCAATCAATTAATAGAGCAGTTAATAAAAAAGGAAAGGTTCTATCTGGCATGG 4155
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2279 TGCCCGCCCAAGGCGATCGCGCGCAACGACGAGATCGACAGCTGGTGAGCAAGGGCA 2338
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4156 TACGACACAAAGGAATTTGGAGGAAATGAACAGTAGATAATATTAGTCAGTGTCTGGAA 4215
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2339 TCAGCAAGTGTCTTCTCGGAGCGGATCAT 2370
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4216 TCAGGAAGTACTATTTTAGATGGATAGAT 4247
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 11

```
US-09-943-286-3
; Sequence 3, Application US/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02.UT
; CURRENT APPLICATION NUMBER: US/09/943,286
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 8933
; TYPE: RNA
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)...(8933)
; OTHER INFORMATION: Sequence of transcripts produced from the BH10
; OTHER INFORMATION: plasmid.
US-09-943-286-3
```

```
Query Match 43.7%; Score 1079.6; DB 10; Length 8933;
Best Local Similarity 55.7%; Pred. No. 1.7e-170;
Matches 1341; Conservative 281; Mismatches 734; Indels 52; Gaps 5;

QY 14 TGSCCGAGGCGCATGAGCCAGG---CCACCAGCGCCACATCTCTGATGCAGGCGAGCAACT 70
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1200 UGCGUGAAGCAUGAGCCAAAGUACAAAUACAGCUACAAUAAUGAUGCAGAGAGGCAAU 1259
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGTCTTCAACTCGGCAAGGAGGCGGCACATCCGCC 130
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1260 UHAGGAACCAAGAAAGAUUGUUAAGUUUAUUGGCAAGAAAGGCGCACACACGCCA 1319
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 131 GCAACTGCGCGCGCCCGCCCAAGAAAGGCTCTGGAAGTGCAGCAAGAGGCGCCACACAGA 190
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1320 GAAAUUGCAGGGCGCCUAGGAAAAAGGCGUUGUAAAAUUGGAAAGGAAAGGACACCAAA 1379
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 191 TGAAGGACTGCACGAGCGCCAGGCCAACTTCTTCCCGGAGGAGCTGGCTTCCCGCCAGG 250
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1380 UGAAAGAUUGUACUGAGAGACAGGCUAA-UUUUUUAGGGAAGAUUCUGGCCUUCUACAA 1438
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 251 GCAAGGCGCCCGAGT-----TCCCCAGCG 274
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1439 GGAAGCGCCAGGAUUUUUUCAGAGCAGACAGCCACAGCCCCACCAUUUCUUCAG 1498
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 275 AGCAGAACCCCGCCCAACAGCCCGCCAGCTGCAGTGCAGTGCAGCG-----CGACA 328
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1499 AGCAGACAGAGCCCAACAGCCCGCCAGAGAGCUUAGGUCUGGGUAGAGACAACA 1558
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 329 ACCCGCGCAGGCGCGCGCGCGCGCGCGCGCA-----CCCTGAACCTTCCCCCAGA 382
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1559 ACUCCCGCCUAGAGCAGGAGCGCGGAGCAAGCAAGCAUCCUUAUCCUCCUAGA 1618
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 383 TCACCTCTGTGCGAGCGCCCTGTGTAGCATCAAGTGGCGCGCCAGATCAAGAGAGGCC 442
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1619 UCACUUCUUGGCAACGACCCUUGUCACAAUAAAGAUAGGGGGCAACUAAAGAAAGCUC 1678
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 443 TGTGTGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 502
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1679 UAUUAGAUACAGGAGCAGAUACAGAUUUAAGAAAUAGAUUUGCCAGGAAGAU 1738
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 503 AGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCAAGTGGCGCGCGCGCGCGCGCGCG 562
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1739 AACCAAAAUAGUAGGGGGAUUGGAGUUUAUCAAGAAAGACAGAUUGAUGACAGAUAC 1798
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 563 TGATCGAGATCTCGGCAAGAGCCCATCGCACCGTGTCTGCGCGCGCGCGCGCGCGCGCG 622
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1799 UCAUAGAAUUCUGUGGACAUAAAGCUAAGUACAGAUUUAUAGUAGGACCUACACCUUCA 1858
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 623 ACATCATCGCGCGCAACATCTGACCCAGCTGGGCGTGCACCCCTGAACCTTCCCCATCAGCC 682
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1859 ACAUAAUUGGAAGAAUUCUGUACUCAGAUUGUUGCAUUAUUAUUAUUAUUAUUAUUA 1918
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 683 CCATCAGACCGTGGCGGTGAAGCTGAAGCCCGGCGATGGAGCGGCCCGCAAGGTGAAGCAGT 742
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1919 CUUUGAGACUGUACCAAGUAAAAUUAAGCCAGGAUUGGAGGCCCAAAAAAUAAAAAU 1978
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 743 GGCCCTGTACCGAGGAGAAAGTCAAGCCCTGACCGCCATCTGCGAGGAGATGAGGAAGG 802
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1979 GGCAUUGACAGAGAAAAAUAAGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2038
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 803 AGGGCAAGATCACCAAGATCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 862
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db 1320 GAAUUGCAGGGCCCCUAGGAAAGGGGCUUGUUGAAAUUGUGAAAGGAAAGGACACAAA 1379
Qy 191 TGAGGAGTACCGAGCGCCAGCCAGCAACTTCTTCGCGAGGAGCTGCGCTTCCGCCAGG 250
Db 1380 UGAAGAAGUUGACUGAGAGAGCGCUAA-UUUUUUAGGGAAGAUUGGCUUCCUACAAG 1438
Qy 251 GCAAGGCCCGCGAGT-----TCCCCAGCG 274
Db 1439 GGAAGGCCAGGAUUUUUUCAGAGCAGACAGCCACAGCCCCACCAUUUUCUAG 1498
Qy 275 AGCAAAACCGGGCCAGACAGCCCGAGCGAGCTGCGAGTGCAGGTGCGGG-CCACA 328
Db 1499 AGCAGCAGAGGCAACAGCGCCACCAAGAGAGAGCUUACGUCUGGGUAGAGACAA 1558
Qy 329 ACCCCGAGGAGCGCGCGCGAGCGCCAGGCA-----CCCTCAACTTCCCCAGA 382
Db 1559 ACUCCCCUCAGAAAGAGGAGCGCGAGCAAGGAACUGUAUCCUUUAUCCUCCAG 1618
Qy 383 TCACCTGTGGCAGCGCCCTGTGTAGCTATCAAGTGGGCGGCGAGATCAAGAGAGGCC 442
Db 1619 UCACUUGGCAACGACCCCGUCACAAUAAAGUAGGGGGCAACUAAAGAAAGCUC 1678
Qy 443 TGCTGGACAGCGCGCGAGCAGACCTGTGTGGAGGAGATGAGCTCCCGCGCAAGTGA 502
Db 1679 UAUAGAUACAGGAGCAGAUACAGAUUAGAGAAUAGAUUUGCCAGGAAGUAGA 1738
Qy 503 AGCCCAAGATGATCGCGGATCGCGGCTTTCATCAAGGTGCGCGAGTACGACAGATCC 562
Db 1739 AACCAAAAUAGAGGGGAUUGGAGGUUUUAUCAAAGUAAGACAGAUUAGUACAGAUAC 1798
Qy 563 TGATCGAGATCTGCGGCAAGAGCCATCGCGACCGTGTGATCGGCGCCACCGCCGTGA 622
Db 1799 UCAUAGAAUUCUGUGGACAUAAAGCUUAGGUAUAGUAGGAGGACCUACACCUUCA 1858
Qy 623 ACATCATCGCGCGCAACATGCTGACCGAGCTGGGCTGCACCTGAACTTCCCATGAGCC 682
Db 1859 ACAUAAUUGGAAGAAUUCUGUUGACUAGAUUGGUUGACAUUUAUUUCCCAUAGCC 1918
Qy 683 CCATCGAGACCGTCCCGTGAAGCTGAAGCCGCGCATGAGCGGCCCAAGGTGAAGCAGT 742
Db 1919 CUUUGAGAGUUGUACAGUAAAUUAAAGCAGGAUUGGAGGCGCCAAAGAUAAACAU 1978
Qy 743 GGCCCTTGACGAGGAGAAAGTAAAGCCCTTGACCGCCATCTGTGCGAGGAGATGGAAGG 802
Db 1979 GGCCAUUGACAGAGAAAUAAAGCAUUAUAGAAAUUUUUGACAGAAUUGGAAAGG 2038
Qy 803 AGGCAAGATCACCAAGTCCGCGCCCGAGACCCCTACACACCCCGTGTCCCATCA 862
Db 2039 AAGGAAAUUUCAAAAUUGGCGGCUAGAAUCCAUACAUAUCCAGAUUUUGCCAUAA 2098
Qy 863 AGAAGAGGACAGCACCAAGTGGCGCAAGCTGTGTGGACTTCCGCGAGCTGAACAAGCGCA 922
Db 2099 AGAAAAGACAGUACUUAUUGGAGAAAUUAGUAGAUUUCAGAGAAUUAUAAAGAA 2158
Qy 923 CCCAGGACTTCTGGGAGTGCAGTGGGCATCCCGCCACCGCGCGCTGGAAGAAGAA 982
Db 2159 CUCAAGACUUCUGGGAAGUUAUUAUAGAAUACCAUCCGCGAGGUUAAAAAGAAA 2218
Qy 983 AGAGCGTACCGTGTGAGCTGGCGGACGCTTACTCAGGTGCGCTGACGAGGACT 1042
Db 2219 AAUCAGAAACAGUAGUUGGUGUAGUAGUUAUUUUUAGUCCUUAUAGAAAGAU 2278
Qy 1043 TCCCAAGTACACCGCTTCCACCATCCCGAGCATCAACAGAGAGCCCGCGCATCGCT 1102
Db 2279 UCAGGAAGUACUGCAUUUACCAUACCUAGUAUAAACAUAGACACACCGGAUAGAU 2338
Qy 1103 ACCAGTACAAGCTGCTGCCCGAGGCTGGAAGGAGCGCCCGAGCATCTTCCAGAGCAGCA 1162
Db 2339 AUCAGUACAAUUGUCCUACAGGGAUGGAAGGAUACACAGCAAAUUAUCCAAAGUAGCA 2398
Qy 1163 TGACAGATCTTGAGCGCTTCCCGCGCGCAACCGCGAGATCGTGTATCTACAGTACA 1222
Db 2399 UGACAAAUAUAGAGCCUU 2458

Qy 1223 TGGACGACCTTGTAGCTGGCGAGCAGCACCTGGAGATCGCCAGCAGCCGCCAAGATCGAGG 1282
Db 2459 UGGAUUAUUUGUAGUAGGACUAGUAGAAUAGGAGCAUAGAAACAAAUAUAGAGG 2518
Qy 1283 AGTGGCCAGACACCTGCTGCTGGTGGGCTTACCAACCCCGGAGCAAGAACAGCAGAGG 1342
Db 2519 AGCUGAGACAAUUCUUGUUGGUGGGAUUUACCAACACAGACAAAACAAUAGAAAG 2578
Qy 1343 AGCCCCCTTCTGTGATGGCTACGAGTGCACCCGAGCAAGTGGAGCTGCAGCCCA 1402
Db 2579 AACUCCAUUCCUUGUUGGUGUUAUAAUCCUUAUAAUAAUAGGACAGUAGAGCCUA 2638
Qy 1403 TCGAGCTGCCCGAGAGAGAGTGGACCGTGAACGATCTCCAGAACTGCTGGCGCAAGC 1462
Db 2639 UAGUGUCGCCAGAAAAGACAGUCUGACUGUAUAGUAGUAGUAGUAGUAGUAGUAG 2698
Qy 1463 TGAATGGCGCAGCAGATCTACCCCGGATCAAGTGGCGGAGCTGTGCAAGTGTGTCG 1522
Db 2699 UGAUUGGGCAAGUAGAUUACCCAGGAAUUAAGAAAGCAUUAUAGUAAUCCUUA 2758
Qy 1523 GCGCGCCAAAGGCTGACCGACATCTGCTGCCCTGACCGAGGAGCGAGCTGGAGCTGG 1582
Db 2759 GAGGAACCAAGACACUACAGAAUAAUACCAUACAGAAAGAGCAGAGUAGAAUUG 2818
Qy 1583 CCGAGAACCGCGAGATCTGCGCGAGCCGCTGCACGGCTGTACTAGCAACCCAGCAAGG 1642
Db 2819 CAGAAAACAGAGACAUUUAUAAAGAACACAGUAGUAGUAGUAGUAGUAGUAGUAG 2878
Qy 1643 ACCTGTGGCGGAGATCCAGAGCAGGCGCAGCAGCTGAGTGGACCTTACAGATCTAC 1702
Db 2879 ACUUAUAGAGAAUUAACAGAGCAGGCGCAAGGCCAAUGGACAUUACAAUUAUUA 2938
Qy 1703 AGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAGATCGCGACCGCCACACCAACG 1762
Db 2939 AGCCAUUAAAUUUCUUAUAAAGAAACAGAAAUUAUAGCAUAGAGGCGCCACACUA 2998
Qy 1763 ACGTGAAGCAGCTGACCGAGGCGCTGCAGAAAGATCGCCATGGAGAGCATCTGTG 1822
Db 2999 AUGUAAACAAUUAACAGAGCGCAGUGCAAAAUAAUAAACACAGAAAGUAGUAG 3058
Qy 1823 GCAAGACCCCAAGTTCGCTGCGCTGCCATCCAGAGGAGACCTGGGAGACCTTGTGG 1882
Db 3059 GAAAGACUCCUAAAUUUAUAAUACUACAAAAGAAACAAUAGGAAACUAGGUGGAG 3118
Qy 1883 ACTACTGGCGGCGCACCTGGATCCCGAGTGGGAGTTCGTAACACCCCGCTGTGTA 1942
Db 3119 AGUUAUGGCAAGCCACUGGAUUCUGAGUGGAGUUAUUAUAAUAAUAAUUAUAG 3178
Qy 1943 AGCTGTGTACAGCTGGAGAGGAGCCATCATCGGCGCGGAGACCTTCTACGTGGAG 2002
Db 3179 AAUUAUGGUACCAUAGAGAAAGAACCCUAGUAGGAGCAGAAACCUUUAUAGUAG 3238
Qy 2003 GCGCGCCCAACCGGAGCAGCAAGATCGGCAAGCGCGCTAGTGCACCGCGGCGCG 2062
Db 3239 GGGCAGCUAACAGGAGACUAAAUUAGGAAAGCAGGAUUAUUAUAAUAAUAAUAA 3298
Qy 2063 AGAAGATCTGAGCTGACCGAGACCAACACAGAGACCGAGCTGCAGCGCATCCAGC 2122
Db 3299 AAAAGGUUUGCCUUAACUACACACAAUACAGAAACUAGUUAUAAUAAUUAU 3358
Qy 2123 TGGCCCTCAGGAGCAGGCGGAGCGAGGTGAACATCGTGACCGCAGCAGCTACGCTGG 2182
Db 3359 UAGCUUUGCAGGAUUCAGGAUUAAGAAUAAUAAUAAUAAUAAUAAUAAUAAUAA 3418
Qy 2183 GCATCATCCAGCGCCCGGAGCAGAGCGAGCGAGCTGTAACAGATCATCTGAGC 2242
Db 3419 GAUUAUUAUAGCAGACACCAUAAUAAUAAUAAUAAUAAUAAUAAUAAUAAUAA 3478
Qy 2243 AGCTGATCAAGAAAGAGAGGTGTACCTGAGCTGGGTGCCCGCCAGGCGCATCGCG 2302
Db 3479 AGUUAUAAUAAAGGAAAGGCUUAUUGGCAUGGAGUAGGACACACAAAGGAAUUGG 3538

[illegible]

RESULT 14

[illegible]

LOCATION: (106)..(1641)
OTHER INFORMATION: encodes proteins of viral inner core
NAME/KEY: misc.feature
LOCATION: (1401)..(3617)
OTHER INFORMATION: encodes viral proteins but not integrase
NAME/KEY: misc.feature
LOCATION: (3708)..(5715)
OTHER INFORMATION: regulates high-level production of HIV genes
NAME/KEY: misc.feature
LOCATION: (3847)..(5944)
OTHER INFORMATION: encodes protein regulating the transfer of RNA to cytoplasm
NAME/KEY: misc.feature
LOCATION: (3939)..(4181)
OTHER INFORMATION: gene participates in viral assembly and budding
NAME/KEY: misc.feature
LOCATION: (4099)..(4941)
OTHER INFORMATION: encodes truncated form of viral coat protein
NAME/KEY: misc.feature
LOCATION: (6098)..(9918)
OTHER INFORMATION: vaccine vector pGAI
US-09-798-675-5

Query Match 43.5%; Score 1073.2; DB 10; Length 9918;
Best Local Similarity 67.2%; Pred. No. 1.9e-169;
Matches 1618; Conservative 0; Mismatches 738; Indels 52; Gaps 5;
QY 14 TGGCCGAGCCATGAGCCAGG---CCACACAGCGCCACATCCTGTGATGCGAGCGCAGCAACT 70
DB 1193 TGGCTGAAGCAATGAGCCAAAGTAAACAAATACAGCTACCAATATGATGCGAGAGGCAATT 1252
QY 71 TCAAGGCCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGCGGCACATGCGCCC 130
DB 1253 TTAGGACCAAGAAAGATGTTTAAGAGCTTCAATAGCGGCAAGAAAGGCGCACACAGCCA 1312
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGCTGTGGAAGTCTGCGGCAAGGAGGCGGCACACAGA 190
DB 1313 GAAATTCAGGGGCCCCCTAGGAAAGAGGCGAGCTGGAAAGCGGAAAGAGAGAGACACAAA 1372
QY 191 TGAAGGACTCACCGAGCGCGCGAGCAACTCTTCGCGAGGAGGAGGCTGCGGCTTCCCGCAGG 250
DB 1373 TGAAGAGTTGACTGAGAGACAGGCTAA-TTTTGTAGGAGATCTGGGCTTCCCTACAG 1431
QY 251 GCAAGGCGCCGAGT-----TCCCGCAGCG 274
DB 1432 GGAAGGCGAGGAATTTCTTTCAGAGCAGACAGCCACAGCCACCACTTCTTCAG 1491
QY 275 AGCAGACCGCGCCACAGCCCGCCAGCGCGAGCTGCGAGGTGCGCGG-----CGACA 328
DB 1492 AGCAGACCGAGCCCAAGCGCCCGCCAGAGAGAGCTTCAGGCTGCGGAGTAGAGACAA 1551
QY 329 ACCCGCGCAGCGCGCGCGCGAGCGCCAGGCGCA-----CCCTGAAGCTTCCCGCAGA 382
DB 1552 ACTCCCGCTCAGAAGCAGGAGCGGATAGACAAGAACTGTATCTTTAACTTCCCTCAGA 1611
QY 383 TCACCTGTGGAGCGCCCGCTGTGAGCATCAAGTGGCGGCGGCAGATCAAGAGGCGCC 442
DB 1612 TCACCTTTGGCAGCAGCCCTGCTCAACAATAAGATAGGCGGCGCACTAAAGGAAGCTC 1671
QY 443 TGGTGGACCGCGCGCGAGCACCGCTGTGGAGGAGATGAGCTGCGCGGCAAGTGA 502
DB 1672 TATTAGATACAGGAGCAGATGATACAGTATTAGAGAAATGAGTTTCCAGGAAGATGA 1731
QY 503 AGCCCAAGATGATGCGCGGCGATCGCGGCTTCATCAAGGTGCGCGAGTAGACACAGATCC 562
DB 1732 AACCAAAATGATAGGCGGAATTGGAGGTTTATCAAAAGTAAGACACTATGATCAGATAC 1791
QY 563 TGATCGAGATCTGGGCGAAGAGCCATCGCGCCGCTGTGATGCGGCGCCACCGCCCGTGA 622
DB 1792 TCATAGAAATCTGTGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGCA 1851
QY 623 ACATCTTCGCGCGCAACATGCTGACCCAGCTGGGCTGCACCCCTCAACTTCCCGCATCAGC 682
DB 1852 ACATAATTGGAAGAAATCTGTTGACTCAGATGTTGCTGCACCTTTAAATTTTCCCATTAGCC 1911

QY 683 CCATCGAGACCGTCCCGCTCAAGCTGAAGCCCGCATGAGCGCCCAAGGTGAAGCAGT 742
DB 1912 CTATTCAGAGCTGTACCAGTAATAATTAAGCCAGGAATGATGCGCCCAAAAGTTAAACAAT 1971
QY 743 GGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCCCTCTCGGAGGAGATGAGGAAG 802
DB 1972 GGCCATTGACAGAAGAAATAAAGCAATTAGTAGAAATTTGTACAGAAATGGAAGG 2031
QY 803 AGGGCAAGATCACCAGATCGGCCCGCGAGAACCCCTACACACCCCGCTGTTCCGCCATCA 862
DB 2032 AAGGGAAATTTCAAAATTTGGCCCTGAGAAATCCATACATACTCCAGTATTTGCCATAA 2091
QY 863 AGAAGAGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCCGGAGCTGCAACAAGCGCA 922
DB 2092 AGAAAAAGACAGTACTAATGGAGAAATTTAGTAGATTTTCAGAGACTTTAATAAGAGAA 2151
QY 923 CCCAGGACTTCTGGAGGTGCACTGGGCATGCCCGACCCCGCGGCTGCAAGAAGAGA 982
DB 2152 CTCAGAGACTTCTGGGAAGTTCAATTAGGAATAGCACATCCCGAGGGTTAAAAAAGAAA 2211
QY 983 AGAGCGTGACCGTCTGGAGGTGGCGGAGCGCTTCTACGCGTGCCTCGAGGAGACT 1042
DB 2212 AATCAGTAACAGTACTGGATGGGTGATGCATATTTTCAGTTCCTTASATGAAGACT 2271
QY 1043 TCCGCAAGTACACCGCTTCCACCATCCCCAGCATCAACACAGAGACCCCGCGCTCCGCT 1102
DB 2272 TCAGGAGTATATGCAATTTACCATACCTAGTATAACAAATGAGACACCGAGGATTAGAT 2331
QY 1103 ACCAGTACAACTGTGCGCCCGAGGTGGAAGGCGAGCCCGCAGCTTTCAGAGCAGCA 1162
DB 2332 ATCAGTACAAATGTCTTCCACAGGGATGGAAGGATCACCAGCAATATTCTCAAGTAGCA 2391
QY 1163 TGACCAAGATCTGGAGCGCTTCCGCGCCGCAACCCCGAGAGTCTGATCTACCACTACA 1222
DB 2392 TGACAAAAATCTTAGAGCCCTTTAAAAAACAATAATCCAGACATAGTTATCTATCAATACA 2451
QY 1223 TGGACCACTCTAGCTGGCGAGCAGCTGGAGATTCGCGCAGCAGCCCGCGCAAGATCGAGG 1282
DB 2452 TGAACGATTTGTATGATAGGATCTGACTTAGAAATAGGCGAGCATAGAACAAAAATAGAGG 2511
QY 1283 AGCTGCGCAGACCTGCTGCGCTGGGCTTACCACCCCGCAACAAGACCCAGAGG 1342
DB 2512 AGCTGAGACAACTCTGTTGAGGTGGGACTTACCACCGAGCAAAAAACATCAGAAAG 2571
QY 1343 AGCCCGCTTCTGTGATGGGTACAGAGCTGCAACCCCGCAACAGTGAGCGTGCAGGCCA 1402
DB 2572 AACCTCACTTCTGTGATGGGTATGAACTCCATCTGATAAATGACAGTACAGCCTA 2631
QY 1403 TCAGCTGCGCGAGAGGAGCTGGACCGTGAACGACATCCAGAACTGGTGGGCGAAGC 1462
DB 2632 TAGTGTGCGCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAATTTAGTGGGAAAT 2691
QY 1463 TGAACCTGGCGCAGCAGATCTACCCCGCATCAAGGTGCGCGAGCTGTGAAGCTGTGC 1522
DB 2692 TGAATACCGCAAGTCAAGTTTCCAGCGGATTAAGTAAAGCAATTTATGTAACCTCTTA 2751
QY 1523 GCGCGCGCAAGGCGCTGACCGACATCTGCGCCCTTGACCGAGGAGCGCGAGCTGGAGTGG 1582
DB 2752 GAGSAACCAAGCAGTAAACAGAAATTAACACTAACAGAAAGAGCAGAGCTAGAACTGG 2811
QY 1583 CCGAAGAACCGGAGATCCTGCGCGAGCGGTGACGCGGTGTACTACGACCCCGAGCAAGG 1642
DB 2812 CAGAAAAACAGAGAGATTTCTAAAAAGAACCGTACATGAGGTGATTTATGACCCATCAAAAG 2871
QY 1643 ACCTGGTGGCGGAGATCCAGAAAGCAGGCGCACGACAGTGGAGCTTACCAGATCTACCAAG 1702
DB 2872 ACTTAATAGCAGAAATACAGAGCAGGCGGAGCCCAATGACATATCAAAATTTATCAAG 2931
QY 1703 AGCCCTTCAAGAACTGGAAGCCGCGCAAGTACGCCAAGATGCGCACCGCCCGCACCAACG 1762
DB 2932 AGCCATTTAAAAATCTGAAACACAGAAAAATATGCAAGAAATGAGGGGTGCCCCACACTAATG 2991

QY	1763	ACGTGAAGCAGCTGACCGAGCCGTGCAGAGATCGCATGGAGAGCATCGTGAICTGGG	1822
Db	2992	ATGTAAACAATTAACAGAGCGAGTGCATAAAATAAACACAGAAGCATAGTAATATGGG	3051
QY	1823	GCAAGACCCCCAAGTTCCGCTGCCATTCAGAAGGAGACCTGGAGACCTGGTGACCG	1882
Db	3052	GAAAGACTCTAAATTTAACTACCCATACAAAGAGAAACATGGGAACATGGTGACAG	3111
QY	1883	ACTACTGGCAGGCCACCTGGGATCCCCGAGTGGGAGTTCTGTAACACACCCGCCCTGGTGA	1942
Db	3112	AGTATTGGCAAGCCACCTGGGATTCCTGAGTGGGAGTTTGTAAATACCCCTTCCTTAAGTGA	3171
QY	1943	AGCTGTGTTACCACTGTCGAGAGGAGCCCATCATCGCGCCGAGACCTTCTACGTGGAGC	2002
Db	3172	AAATTATGGTACCAGTTAGAGAAAGAACCCATAGTAGGACGAGAAACCTTCTATGTAGATG	3231
QY	2003	GGCGCCCAACCGCAGAGACCAAGATCGCAAGCGCGCTACGTGACCGACCGGGCGCGC	2062
Db	3232	GGCAGCTAACAGGAGAGACTAAATTAGGAAAGCAGGATATGTTACTACAAAGGAAGAC	3291
QY	2063	AGAAGATCGTGAAGCTTGACCGAGACCAACCAAGACCGAGCTGCAGGCCCATCCAGC	2122
Db	3292	AAAAGGTTGTCCCCCTAACTAACCAACAATAACAGAAATCAAGTACAAAGCAATTTATC	3351
QY	2123	TGGCCCTCGAGGACAGCGCAGCGAGGTGAACATCTGTACCGACAGCCAGTACGCCCTGG	2182
Db	3352	TAGCTTTGCAGGATTCAGGATTTAGAAGTAACATAGTAACAGACTCACATATGCAATTAG	3411
QY	2183	GCATCATCCAGCCOAGCCGACAGAGCGAGCGAGCTGGTGAACAGACTCATCGAGC	2242
Db	3412	GAATCATTTCAAGCAACACAGATAAAAGTGAATCAGAGTTAGTCAATCAATAATAGAGC	3471
QY	2243	AGCTGATCAAGAGGAGAGGTGTACTGTAGCTGGGTGCCGCCCAACAAGGGCATCGGGC	2302
Db	3472	AGTTAATAAAAAGGAAAGGTTCTATCTGGCATGGGTACCAGCACACAAAGGAATGGAG	3531
QY	2303	GCAACGACGAGATGCACAAGCTGTTGAGCAAGGSCATCCGCAAGGTCGTCTCCCTGGAGC	2362
Db	3532	GAAATGACACAGTAGATAAATTAGTCACTGCTGGAATCAGGAANAATCTATTTTAGATG	3591
QY	2363	GCATCGAT	2370
Db	3592	GAATAGAT	3599

Search completed: February 10, 2003, 21:10:57
job time : 153.929 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 10:53:38 : Search time 6249 Seconds
(without alignments) 11470.669 Million cell updates/sec

Title: US-09-610-313-31
Perfect score: 2463
Sequence: 1 gtcagcgcacccatggccga.....gggctagcaccggtaattc 2463

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1:	gb_ba.*	GenEmbl.*
2:	gb_htg.*	
3:	gb_in.*	
4:	gb_om.*	
5:	gb_ov.*	
6:	gb_pat.*	
7:	gb_ph.*	
8:	gb_pl.*	
9:	gb_pr.*	
10:	gb_ro.*	
11:	gb_sts.*	
12:	gb_sy.*	
13:	gb_un.*	
14:	gb_vl.*	
15:	em_ba.*	
16:	em_fun.*	
17:	em_hum.*	
18:	em_in.*	
19:	em_mu.*	
20:	em_om.*	
21:	em_or.*	
22:	em_ov.*	
23:	em_pat.*	
24:	em_ph.*	
25:	em_pl.*	
26:	em_ro.*	
27:	em_sts.*	
28:	em_un.*	
29:	em_vl.*	
30:	em_htg_hum.*	
31:	em_htg_inv.*	
32:	em_htg_other.*	
33:	em_htg_mus.*	
34:	em_htg_pln.*	
35:	em_htg_rod.*	
36:	em_htg_mam.*	
37:	em_htg_vrt.*	
38:	em_sy.*	
39:	em_htgo_hum.*	
40:	em_htgo_mus.*	
41:	em_htgo_other.*	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2463	100.0	2463	6	AX455915 Sequence
2	2442.2	99.2	2469	6	AX455914 Sequence
3	2436.2	98.9	2457	6	AX455916 Sequence
4	2027.6	82.3	9166	6	AX427930 Sequence
5	2019.2	82.0	9788	6	AX427936 Sequence
6	2005.2	81.4	9169	6	AX427931 Sequence
7	2001.8	81.3	9194	6	AX427926 Sequence
8	2000.2	81.2	9194	6	AX427925 Sequence
9	2000.2	81.2	12411	6	AX427927 Sequence
10	1996.8	81.1	9785	6	AX427938 Sequence
11	1993.4	80.9	9170	6	AX427928 Sequence
12	1993.4	80.9	9782	6	AX427935 Sequence
13	1993.4	80.9	9783	6	AX427934 Sequence
14	1993.4	80.9	9792	6	AX427932 Sequence
15	1991.8	80.9	9189	6	AX427921 Sequence
16	1988.8	80.7	4353	12	AF287353 Synthetic
17	1972.2	80.1	3015	6	AX455946 Sequence
18	1955.6	79.4	3009	6	AX455987 Sequence
19	1886.6	76.6	2577	6	AX457088 Sequence
20	1879	76.3	4343	6	AX306428 Sequence
21	1875	76.1	4341	6	AX306429 Sequence
22	1865.4	75.7	1977	6	AX455954 Sequence
23	1857.8	75.4	1965	6	AX455952 Sequence
24	1852.4	75.2	7897	6	AX427923 Sequence
25	1798.8	73.0	3024	12	AF287355 Synthetic
26	1784	72.4	9407	6	AX427937 Sequence
27	1651.8	67.1	4307	6	AX019132 Sequence
28	1632.6	66.3	4307	6	AX035453 Sequence
29	1632.6	66.0	4307	6	AX056836 Sequence
30	1624.6	66.0	4307	6	AX283597 Sequence
31	1624.6	66.0	4307	6	AX283607 Sequence
32	1624.6	66.0	4327	6	AX035465 Sequence
33	1624.6	66.0	4327	6	AX035465 Sequence
34	1624.6	66.0	4642	6	AX035463 Sequence
35	1624.6	66.0	4642	6	AX035463 Sequence
36	1624.6	66.0	9772	6	AX283584 Sequence
37	1581.2	64.2	1680	6	AX455960 Sequence
38	1575.2	64.0	1688	6	AX455962 Sequence
39	1530.6	62.1	8366	6	AX180752 Sequence
40	1518.2	61.6	3981	6	AX306430 Sequence
41	1300.8	52.8	4288	6	AX149648 Sequence
42	1255.8	51.0	9010	14	AF110974
43	1255.8	51.0	9010	14	AF110975

ALIGNMENTS

RESULT 1	AX455915	Sequence 31 from Patent WO0204493.	2463 bp	DNA	linear	PAT 06-JUL-2002
LOCUS	AX455915					
DEFINITION	Sequence 31 from Patent WO0204493.					
ACCESSION	AX455915					
VERSION	AX455915.1	GI:21714900				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						

CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

Location/Qualifiers

Source

1. .2463

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="PR975YM"

BASE COUNT 567 a 835 c 759 g 302 t

ORIGIN

Query Match 100.0%; Score 2463; DB 6; Length 2463;

Best Local Similarity 100.0%; Pred. No. 1.9e-249;

Matches 2463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACGCCACCATGGCCGAGGCGCATGAGCCAGCCAGCCAGCCACATCTCTGATCCAG 60
DB 1 GTGACGCCACCATGGCCGAGGCGCATGAGCCAGCCAGCCAGCCACATCTCTGATCCAG 60
QY 61 CGCAGCAACTTCAAGGCCGCCAAGCGCATCATCAAGTGTCTTCAACTCGCGCAAGAGGGC 120
DB 61 CGCAGCAACTTCAAGGCCGCCAAGCGCATCATCAAGTGTCTTCAACTCGCGCAAGAGGGC 120
QY 121 CATATCGCCGCAACTCGCGGCCGCCGCCGCCAGAGGGCTGTGGAAGTGGGGCAAGGAG 180
DB 121 CATATCGCCGCAACTCGCGGCCGCCGCCGCCAGAGGGCTGTGGAAGTGGGGCAAGGAG 180
QY 181 GGCACCCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCC 240
DB 181 GGCACCCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCC 240
QY 241 TTCCGCCAGGGCAAGGCCCGCGAGTTCCCGAGCGAGAGAACCGCGCAACAGCCCCACC 300
DB 241 TTCCGCCAGGGCAAGGCCCGCGAGTTCCCGAGCGAGAGAACCGCGCAACAGCCCCACC 300
QY 301 AGCGCGAGCTGAGGTGCGCGGGCGCAACCCCGCAGCGAGCGCGCGCGCGCGCGCAG 360
DB 301 AGCGCGAGCTGAGGTGCGCGGGCGCAACCCCGCAGCGAGCGCGCGCGCGCGCGCAG 360
QY 361 GGCACCTGAACTTCCCGCCAGATCAACCTGTGGCAGCGCCCGCTGTGTGAGCATCAAGTG 420
DB 361 GGCACCTGAACTTCCCGCCAGATCAACCTGTGGCAGCGCCCGCTGTGTGAGCATCAAGTG 420
QY 421 GCGCGCAGATCAAGAGGGCCCTGCTGGACACCGCGCGCGAGACACCGTGTGTGGAGGAG 480
DB 421 GCGCGCAGATCAAGAGGGCCCTGCTGGACACCGCGCGCGAGACACCGTGTGTGGAGGAG 480
QY 481 ATGAGCTGCGCGCAAGTGAAGCCCAAGATGATCGCGCATCGCGGCTTTCATCAAG 540
DB 481 ATGAGCTGCGCGCAAGTGAAGCCCAAGATGATCGCGCATCGCGGCTTTCATCAAG 540
QY 541 GTGCGCCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGAGGCGCATCGGCACCTG 600
DB 541 GTGCGCCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGAGGCGCATCGGCACCTG 600
QY 601 CTGATCGGCCCGCCCGCGTGAACATCATCGGCGCGCAACATGCTGACCGAGTGGGTGC 660
DB 601 CTGATCGGCCCGCCCGCGTGAACATCATCGGCGCGCAACATGCTGACCGAGTGGGTGC 660
QY 661 ACCCTGAACTTCCCGCATCAGCCCATCGAGACCGTGGCGCGTGAAGCTGAAGCCCGGATG 720
DB 661 ACCCTGAACTTCCCGCATCAGCCCATCGAGACCGTGGCGCGTGAAGCTGAAGCCCGGATG 720
QY 721 GACGGCCCAAGGTGAAGAGTGGCCCTGACCGAGGAGAGATCAAGGGCCCTGACCGCC 780
DB 721 GACGGCCCAAGGTGAAGAGTGGCCCTGACCGAGGAGAGATCAAGGGCCCTGACCGCC 780
QY 781 ATCTCGGAGGAGTGAAGAGGGGCAAGATCACCAAGATCGCGCGCGCGCGCGCGCTAC 840
DB 781 ATCTCGGAGGAGTGAAGAGGGGCAAGATCACCAAGATCGCGCGCGCGCGCGCTAC 840
QY 841 AACACCCCGGTGTCGGCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGTGGAC 900
DB 841 AACACCCCGGTGTCGGCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGTGGAC 900

QY 901 TTCCGGAGCTGAACAGCGCACCCAGGACTTCTGGAGGTGAGCTGGCATCCCCAC 960
DB 901 TTCCGGAGCTGAACAGCGCACCCAGGACTTCTGGAGGTGAGCTGGCATCCCCAC 960
QY 961 CCCTCGGCGCTGAAGAAGAAGAGAGCGTACCGTGTGTGAGCTGGCGGAGCGCTACTTC 1020
DB 961 CCCTCGGCGCTGAAGAAGAAGAGAGCGTACCGTGTGTGAGCTGGCGGAGCGCTACTTC 1020
QY 1021 AGCGTCCCTCGGAGGAGGACTTCCGCAAGTACACCGCTTACCATCCCCAGCATCAAC 1080
DB 1021 AGCGTCCCTCGGAGGAGGACTTCCGCAAGTACACCGCTTACCATCCCCAGCATCAAC 1080
QY 1081 AAGGAGACCCCGCGCATCCGCTACCACTACAAGCTGTGCCCGCCAGGCTGGAAGGCGAG 1140
DB 1081 AAGGAGACCCCGCGCATCCGCTACCACTACAAGCTGTGCCCGCCAGGCTGGAAGGCGAG 1140
QY 1141 CCAGCATCTTCAGAGCAGATGACCAAGATCTGTGAGGCGCTTCCGCGCGCGCAAGCCC 1200
DB 1141 CCAGCATCTTCAGAGCAGATGACCAAGATCTGTGAGGCGCTTCCGCGCGCGCAAGCCC 1200
QY 1201 GAGATCGTGTATCTACCAAGCGCCCTCTAGTGGCGAGCGACCTGGAGATCGGCGAGCAC 1260
DB 1201 GAGATCGTGTATCTACCAAGCGCCCTCTAGTGGCGAGCGACCTGGAGATCGGCGAGCAC 1260
QY 1261 CGCGCCAAAGATCGAGGAGCTGCGCAAGCACTGTGCTGGGCTTACACACCCCGAC 1320
DB 1261 CGCGCCAAAGATCGAGGAGCTGCGCAAGCACTGTGCTGGGCTTACACACCCCGAC 1320
QY 1321 AAGAGCACCAAGAGGAGCGCCCTTCTGTGATGGCTACGAGCTGCACCCCGACAG 1380
DB 1321 AAGAGCACCAAGAGGAGCGCCCTTCTGTGATGGCTACGAGCTGCACCCCGACAG 1380
QY 1381 TGGACCTGTCAGCCCATCTGAGCTGCGCGAGAAAGAGAGCTGGACCGTGAACAGATCCAG 1440
DB 1381 TGGACCTGTCAGCCCATCTGAGCTGCGCGAGAAAGAGAGCTGGACCGTGAACAGATCCAG 1440
QY 1441 AAGTGTGGGCAAGCTGAACCTGGGCGAGCAGATCTACCCCGGCGATCAAGTGGCGCAG 1500
DB 1441 AAGTGTGGGCAAGCTGAACCTGGGCGAGCAGATCTACCCCGGCGATCAAGTGGCGCAG 1500
QY 1501 CTGTGCAAGCTGTCGCGGCGCGCAAGCGCTGACGACATCTGCGCCCTGACCGAGGAG 1560
DB 1501 CTGTGCAAGCTGTCGCGGCGCGCAAGCGCTGACGACATCTGCGCCCTGACCGAGGAG 1560
QY 1561 GCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTGTGCGGAGCGCCGTGCACGGGTGTAC 1620
DB 1561 GCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTGTGCGGAGCGCCGTGCACGGGTGTAC 1620
QY 1621 TACGACCCCAAGAGGAGCTGGTGGCGAGATCAGAGAGGCGCGACCGAGTGGAC 1680
DB 1621 TACGACCCCAAGAGGAGCTGGTGGCGAGATCAGAGAGGCGCGACCGAGTGGAC 1680
QY 1681 TACGAGATCTACCAAGAGCGCTTCAAGAACTGAAGACCGGCAAGTACGCCAAGATGGC 1740
DB 1681 TACGAGATCTACCAAGAGCGCTTCAAGAACTGAAGACCGGCAAGTACGCCAAGATGGC 1740
QY 1741 ACCGCCACACCAAGCAGCTGAAGCAGTACCGAGGCGCTGCGAGAGATCGCCATGAG 1800
DB 1741 ACCGCCACACCAAGCAGCTGAAGCAGTACCGAGGCGCTGCGAGAGATCGCCATGAG 1800
QY 1801 AGCATGTGTATCTGGGCGAAGACCCCAAGTTCGCGCTGCGCCATCCAGAGAGACCTGG 1860
DB 1801 AGCATGTGTATCTGGGCGAAGACCCCAAGTTCGCGCTGCGCCATCCAGAGAGACCTGG 1860
QY 1861 GAGACCTGTGTGAGCGGAGTGTGCGAGCGCACCTGGATCCCCGAGTGGAGTTCGTGAAC 1920
DB 1861 GAGACCTGTGTGAGCGGAGTGTGCGAGCGCACCTGGATCCCCGAGTGGAGTTCGTGAAC 1920
QY 1921 ACCCGCCCGCTGTGTAAGCTGTGTACCGAGTGGAGAGGCGCCATCATCGGCGCGAG 1980
DB 1921 ACCCGCCCGCTGTGTAAGCTGTGTACCGAGTGGAGAGGCGCCATCATCGGCGCGAG 1980
QY 1981 ACCTTCTACGTGGAGCGCGCGCCCAAGACCGGAGACCAAGATCGGCAAGCGCGCTACGTG 2040

```
Db 1981 ACCTTCTACGTGGACGGCGCCGCAACGGCGAGACCAAGATCGCAAGCGCGGTACGTG 2040
|||||
QY 2041 ACCGACCGGGCGGCGAGAGATCGTGACCGTGAACCGAGACCAACCAAGACCGAG 2100
|||||
Db 2041 ACCGACCGGGCGGCGAGAGATCGTGACCGTGAACCGAGACCAACCAAGACCGAG 2100
|||||
QY 2101 CTGACGCCATCCAGCTGGCCCTGCAGACAGCGGCGAGAGGTGAACATCGTGACCGAC 2160
|||||
Db 2101 CTGACGCCATCCAGCTGGCCCTGCAGACAGCGGCGAGAGGTGAACATCGTGACCGAC 2160
|||||
QY 2161 AGCCAGTACGCCCTGGCGATCATCCAGGCCAGCCCGACAGAGCGAGAGGAGCTGGTG 2220
|||||
Db 2161 AGCCAGTACGCCCTGGCGATCATCCAGGCCAGCCCGACAGAGCGAGAGGAGCTGGTG 2220
|||||
QY 2221 AACACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGGTGCCGCC 2280
|||||
Db 2221 AACACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGGTGCCGCC 2280
|||||
QY 2281 CACAAGGGCATTCGGCGCAACGAGCAGATCGACAAGCTGGTGAAGGATCCGCAAG 2340
|||||
Db 2281 CACAAGGGCATTCGGCGCAACGAGCAGATCGACAAGCTGGTGAAGGATCCGCAAG 2340
|||||
QY 2341 GTGCTGTCTTGGACGCGATCGATGGCGGATCGTGATCTACAGTACATGGAGCAGCTG 2400
|||||
Db 2341 GTGCTGTCTTGGACGCGATCGATGGCGGATCGTGATCTACAGTACATGGAGCAGCTG 2400
|||||
QY 2401 TACGTGGCAGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAA 2460
|||||
Db 2401 TACGTGGCAGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAA 2460
|||||
QY 2461 TTC 2463
|||
Db 2461 TTC 2463
|||

RESULT 2
AX455914
LOCUS AX455914 2469 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 30 from Patent WO0204493.
ACCESSION AX455914
VERSION AX455914.1 GI:21714899
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof
JOURNAL Patent: WO 0204493-A 30 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
Location/Qualifiers
source
1..2469
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="PR975(+)"
BASE COUNT 571 a 833 c 761 g 304 t
ORIGIN

Query Match 99.2%; Score 2442.2; DB 6; Length 2469;
Best Local Similarity 99.6%; Pred. No. 2.9e-247;
Matches 2460; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 1 GTCACGCCACCATGGCCGAGGCATGAGCAGGCCACCGCCACACATCTCTGATCGAG 60
|||||
Db 1 GTCACGCCACCATGGCCGAGGCATGAGCAGGCCACCGCCACACATCTCTGATCGAG 60
|||||
QY 61 CGCAGCAACTTCAAGGCGCCCAAGCGCATCATCAAGTGCCTTCAACTCGCGCAAGGAGGC 120
|||||
Db 61 CGCAGCAACTTCAAGGCGCCCAAGCGCATCATCAAGTGCCTTCAACTCGCGCAAGGAGGC 120
|||||
QY 121 CACATCGCCGCAACTGCGCGGCCGCCGCCGCAAGAGGGCTGCTGGAAGTGGCGAAGGAG 180
|||||
```

```
Db 121 CACATCGCCGCAACTGCGCGGCCGCCGCCGCAAGAGGGCTGCTGGAAGTGGCGAAGGAG 180
|||||
QY 181 GGCACACAGATGAGGAGTGCACCGAGCGCCAGGCGCAACTTCTTCCGCGAGGAGACCTGGCC 240
|||||
Db 181 GGCACACAGATGAGGAGTGCACCGAGCGCCAGGCGCAACTTCTTCCGCGAGGAGACCTGGCC 240
|||||
QY 241 TTCCGCCAGGCGAAGGCGCGGAGTTCCCGAGCGAGCAAGACCGCGCCCAAGAGCCGCCACC 300
|||||
Db 241 TTCCGCCAGGCGAAGGCGCGGAGTTCCCGAGCGAGCAAGACCGCGCCCAAGAGCCGCCACC 300
|||||
QY 301 AGCGCGAGCTGCGAGGTGCGCGGCGACAAACCCCGCAGCGAGCGCGCGCGCGCGCGCCAG 360
|||||
Db 301 AGCGCGAGCTGCGAGGTGCGCGGCGACAAACCCCGCAGCGAGCGCGCGCGCGCGCGCCAG 360
|||||
QY 361 GGCACCTGACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGGTAGCATCAAGGTG 420
|||||
Db 361 GGCACCTGACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGGTAGCATCAAGGTG 420
|||||
QY 421 GGCGCCAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGACACACCGTGTGGAGGAG 480
|||||
Db 421 GGCGCCAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGACACACCGTGTGGAGGAG 480
|||||
QY 481 ATGAGCCTGCGCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCAAG 540
|||||
Db 481 ATGAGCCTGCGCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCAAG 540
|||||
QY 541 GTGCGCAGTACGACCGATCTCTGATCGAGATCGCGCGCATCGCGCGCTTCATCAAG 600
|||||
Db 541 GTGCGCAGTACGACCGATCTCTGATCGAGATCGCGCGCATCGCGCGCTTCATCAAG 600
|||||
QY 601 CTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
|||||
Db 601 CTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
|||||
QY 661 ACCCTGAACCTTCCCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
|||||
Db 661 ACCCTGAACCTTCCCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
|||||
QY 721 GACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
|||||
Db 721 GACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
|||||
QY 781 ATCTGCGAGGAGTGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
|||||
Db 781 ATCTGCGAGGAGTGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
|||||
QY 841 AACACCCCGGTGTCGCGCATCAAGAAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 900
|||||
Db 841 AACACCCCGGTGTCGCGCATCAAGAAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 900
|||||
QY 901 TTCCGCGAGCTGAACAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
|||||
Db 901 TTCCGCGAGCTGAACAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
|||||
QY 961 CCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
|||||
Db 961 CCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
|||||
QY 1021 AGCGTGCCTTGGACGAGGAGTTCGCGCAAGTACACCGCTTCCAGCATCCCGCAGCATCAAC 1080
|||||
Db 1021 AGCGTGCCTTGGACGAGGAGTTCGCGCAAGTACACCGCTTCCAGCATCCCGCAGCATCAAC 1080
|||||
QY 1081 AACGAGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
|||||
Db 1081 AACGAGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
|||||
QY 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTGGAGCGCTTCCGCGCGCGCGCGCGCG 1200
|||||
Db 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTGGAGCGCTTCCGCGCGCGCGCGCGCG 1200
|||||
QY 1201 GAGATCGTGTACTACCA-----GGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1254
|||||
```

Db 1201 GAGATCGTGATCTACAGTACATGAGACACCTGTAGTGGCAGCAGCCTGGAGATCGGC 1260
QY 1255 CAGCACCAGCGGCAAGATCGAGGAGCTGCGCAAGCAGCCTGCTGGCTGGGCTTCACCAACC 1314
Db 1261 CAGCACCAGCGGCAAGATCGAGGAGCTGCGCAAGCAGCCTGCTGGCTGGGCTTCACCAACC 1320
QY 1315 CCGCACAAGAAGCACCAGAGGAGCGCCCTTCTGCTGGATGGCTACAGCTGCAACCC 1374
Db 1321 CCGCACAAGAAGCACCAGAGGAGCGCCCTTCTGCTGGATGGCTACAGCTGCAACCC 1380
QY 1375 GACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACGAC 1434
Db 1381 GACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACGAC 1440
QY 1435 ATCCAGAAGCTGGTGGCAAGCTGAAGTGGGCGGCGCAGCATCTACCCCGGCATCAAGGTG 1494
Db 1441 ATCCAGAAGCTGGTGGCAAGCTGAAGTGGGCGGCGCAGCATCTACCCCGGCATCAAGGTG 1500
QY 1495 CGCAGCTGTGAAGCTGTGCGGCGGCGCAGGCGCTGACCCAGCATCTGTCGCCCTGACC 1554
Db 1501 CGCAGCTGTGAAGCTGTGCGGCGGCGCAGGCGCTGACCCAGCATCTGTCGCCCTGACC 1560
QY 1555 GAGGAGCGGAGCTGGAGCTGGCCGAGAGACCGGAGATCTGCGCGAGCGCGTGCACGGC 1614
Db 1561 GAGGAGCGGAGCTGGAGCTGGCCGAGAGACCGGAGATCTGCGCGAGCGCGTGCACGGC 1620
QY 1615 GTGTACTAGACCCAGCAGGAGCCTGTGGCGGCGCAGGATCCAGAGCAGGCGGCGCAGGAG 1674
Db 1621 GTGTACTAGACCCAGCAGGAGCCTGTGGCGGCGCAGGATCCAGAGCAGGCGGCGCAGGAG 1680
QY 1675 TGACCTTACAGATCTACAGGAGCGCTTCAAGAACCTGAAGACCGGCGCAAGTACGCGCAAG 1734
Db 1681 TGACCTTACAGATCTACAGGAGCGCTTCAAGAACCTGAAGACCGGCGCAAGTACGCGCAAG 1740
QY 1735 ATGCGCAGCGCCACACCAACGAGCTGAAGAGCTGACCGAGCGCTGCGAGAGATCGCC 1794
Db 1741 ATGCGCAGCGCCACACCAACGAGCTGAAGAGCTGACCGAGCGCTGCGAGAGATCGCC 1800
QY 1795 ATGCGCAGCGCTGTGATCTGGGCGCAGACCCCAAGTTCGCGCTGCCATCCAGAGGAG 1854
Db 1801 ATGCGCAGCGCTGTGATCTGGGCGCAGACCCCAAGTTCGCGCTGCCATCCAGAGGAG 1860
QY 1855 ACCTGGAGACCTGTGGACCGCTACTGGCAGCGCCACCTGGATCCCGAGTGGGAGTTC 1914
Db 1861 ACCTGGAGACCTGTGGACCGCTACTGGCAGCGCCACCTGGATCCCGAGTGGGAGTTC 1920
QY 1915 GTGAACACCCCGCCCTGTGTGAAGCTGTGTACAGCTGAGAGAGCGCCATCATCGGC 1974
Db 1921 GTGAACACCCCGCCCTGTGTGAAGCTGTGTACAGCTGAGAGAGCGCCATCATCGGC 1980
QY 1975 GCGGAGACCTTCTACGTGGACGCGCGCCGCAACCGGAGACCAAGATCGGCAAGCGCGC 2034
Db 1981 GCGGAGACCTTCTACGTGGACGCGCGCCGCAACCGGAGACCAAGATCGGCAAGCGCGC 2040
QY 2035 TAGCTGACCGACCGGCGCGGCGCAGAGATCTGTGAGCTGTACCGAGACCAACCAAGAG 2094
Db 2041 TAGCTGACCGACCGGCGCGGCGCAGAGATCTGTGAGCTGTACCGAGACCAACCAAGAG 2100
QY 2095 ACCGAGCTGACGGCCATTCAGCTGGCCCTGCGAGGACGCGGCGAGCGTGAACATCTGTG 2154
Db 2101 ACCGAGCTGACGGCCATTCAGCTGGCCCTGCGAGGACGCGGCGAGCGTGAACATCTGTG 2160
QY 2155 ACCGAGCGCAGTACCGCTTGGGCATCATCTCCAGGCGCCAGCCGACAGAGCGAGCGAG 2214
Db 2161 ACCGAGCGCAGTACCGCTTGGGCATCATCTCCAGGCGCCAGCCGACAGAGCGAGCGAG 2220
QY 2215 CTGGTGAACCCAGATCATCGAGCAGCTGTATCAAGAGGAGAGGTGTACCTGAGCTGGGTG 2274
Db 2221 CTGGTGAACCCAGATCATCGAGCAGCTGTATCAAGAGGAGAGGTGTACCTGAGCTGGGTG 2280
QY 2275 CCGCGCCACAGGGGATCGGGCGCAACGAGCAGATGACAGCTGTGTGAGCAGGCGCATC 2334
Db 2281 CCGCGCCACAGGGGATCGGGCGCAACGAGCAGATGACAGCTGTGTGAGCAGGCGCATC 2340

QY 2335 CGCAAGTGTCTTCTGTGACGGCATCGTATGGCGGATCTACCAAGTACATGGAC 2394
Db 2341 CGCAAGTGTCTTCTGTGACGGCATCGTATGGCGGATCTACCAAGTACATGGAC 2400
QY 2395 GACCTGTAGCTGGGCGAGCGGCGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACC 2454
Db 2401 GACCTGTAGCTGGGCGAGCGGCGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACC 2460
QY 2455 GGTGAATTC 2463
Db 2461 GGTGAATTC 2469
RESULT 3
AX455916
LOCUS
DEFINITION Sequence 32 from Patent WO0204493.
ACCESSION AX455916
VERSION AX455916.1 GI:21714901
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 artificial sequences.
AUTHORS zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL polypeptides and uses thereof
Patent: WO 0204493-A 32 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
Location/Qualifiers
1..2457
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="PR975YMMW"
BASE COUNT 566 a 837 c 754 g 300 t
ORIGIN
Query Match 98.9%; Score 2436.2; DB 6; Length 2457;
Best Local Similarity 99.6%; Pred. No. 1.2e-246;
Matches 2454; Conservative 0; Mismatches 3; Indels 6; Gaps 1;
QY 1 GTGCGACCGCACCATGTGCGGAGGCGCATGAGCGGCGCCACAGCGCCCAACATCTCTGATGCAG 60
Db 1 GTGCGACCGCACCATGTGCGGAGGCGCATGAGCGGCGCCACAGCGCCCAACATCTCTGATGCAG 60
QY 61 CGCAGCAACTTCAAGGGCGCCCAAGCGCATCATCAAGTGTCTTCAACTGCGGCAAGAGGGC 120
Db 61 CGCAGCAACTTCAAGGGCGCCCAAGCGCATCATCAAGTGTCTTCAACTGCGGCAAGAGGGC 120
QY 121 CACATCGCGCGCAACTGCGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTCCGCGCAAGGAG 180
Db 121 CACATCGCGCGCAACTGCGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTCCGCGCAAGGAG 180
QY 181 GGCACACAGATGAAGGACTGCAACCGAGCGCGCCAGCCCAACTTCTTCCGCGAGGACCTGGCC 240
Db 181 GGCACACAGATGAAGGACTGCAACCGAGCGCGCCAGCCCAACTTCTTCCGCGAGGACCTGGCC 240
QY 241 TTCCCCCAGGGCAAGCGCGCGAGTTCCCGAGCGAGCAACCGCGCCCAACACAGCCCCACC 300
Db 241 TTCCCCCAGGGCAAGCGCGCGAGTTCCCGAGCGAGCAACCGCGCCCAACACAGCCCCACC 300
QY 301 AGCGCGAGCTGCAGTGGCGGCGGACACACCCCGCAGGAGCGCGCGCGCGCGCGCGAG 360
Db 301 AGCGCGAGCTGCAGTGGCGGCGGACACACCCCGCAGGAGCGCGCGCGCGCGCGCGAG 360
QY 361 GGCACCTTGAACCTTCCCGCAGATCACCTGTGTGAGCGCGCCCTGTGTGAGCATCAAGGTG 420
Db 361 GGCACCTTGAACCTTCCCGCAGATCACCTGTGTGAGCGCGCCCTGTGTGAGCATCAAGGTG 420
QY 421 GCGGCGCAGATCAAGGAGGCGCTGTGTGAGACACCGCGCGCGCGCGCGCGCGCGAGGAG 480
Db 421 GCGGCGCAGATCAAGGAGGCGCTGTGTGAGACACCGCGCGCGCGCGCGCGCGCGAGGAG 480

QY 481 ATGAGCCTGCCGGCAAGTGGAGCCCAAGATGATCGCGGCATCGGCGCTTCATCAAG 540
DB 481 ATGAGCCTGCCGGCAAGTGGAGCCCAAGATGATCGGCGGCATCGGCGCTTCATCAAG 540
QY 541 GTGCGCAGTACCAAGCAGATCCCTGATCGAGATCTGCGGCAAGAGGCAATCGGCACCGTG 600
DB 541 GTGCGCAGTACCAAGCAGATCCCTGATCGAGATCTGCGGCAAGAGGCAATCGGCACCGTG 600
QY 601 CTGATCGGCCCAACCCCGTGAACATCATCTCGGCCGAACATGCTGACCCAGCTGGGCTGC 660
DB 601 CTGATCGGCCCAACCCCGTGAACATCATCTCGGCCGAACATGCTGACCCAGCTGGGCTGC 660
QY 661 ACCCTGAACCTTCCCATCATAGCCCATCTGAGACGCTGCCGTGAAGCTGAAGCCCGGCATG 720
DB 661 ACCCTGAACCTTCCCATCATAGCCCATCTGAGACGCTGCCGTGAAGCTGAAGCCCGGCATG 720
QY 721 GACGGCCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAAATCATCAAGGCCCTGACCGCC 780
DB 721 GACGGCCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAAATCATCAAGGCCCTGACCGCC 780
QY 781 ATCTGGAGAGATGAGAGAGGAGGCAAGATCACCAGATCGGCCCGCCGAGAACCCCTAC 840
DB 781 ATCTGGAGAGATGAGAGAGGAGGCAAGATCACCAGATCGGCCCGCCGAGAACCCCTAC 840
QY 841 AACACCCCGCTGTCGCCATCAAGAAAGAGAGCAGCACCAGTGGCGCAGCTGGTGAC 900
DB 841 AACACCCCGCTGTCGCCATCAAGAAAGAGAGCAGCACCAGTGGCGCAGCTGGTGAC 900
QY 901 TTCGGGAGCTGAACAGCGCACCCAGGACTTCTGGGAGTGGAGCTGGGCGATCCCCAC 960
DB 901 TTCGGGAGCTGAACAGCGCACCCAGGACTTCTGGGAGTGGAGCTGGGCGATCCCCAC 960
QY 961 CCGCGCGCTGAAGAAAGAGAGCGTGACCGTGCTGGACGTGGGCGAGCGCTACTTC 1020
DB 961 CCGCGCGCTGAAGAAAGAGAGCGTGACCGTGCTGGACGTGGGCGAGCGCTACTTC 1020
QY 1021 AGCGTGCCCTGGAGGAGGACTTCCGAAGTACACGCCCTTACCATCCCGCAGCATCAAC 1080
DB 1021 AGCGTGCCCTGGAGGAGGACTTCCGAAGTACACGCCCTTACCATCCCGCAGCATCAAC 1080
QY 1081 AACGAGACCCCGGCATCCCTACAGTACAGTGTGTGCCCGCAGGCTGGAGGGGAGC 1140
DB 1081 AACGAGACCCCGGCATCCCTACAGTACAGTGTGTGCCCGCAGGCTGGAGGGGAGC 1140
QY 1141 CCAGCATCTTCAGAGCAGATGACCAAGATCTGTGAGCCCTTCCCGCGCCGCAACCCC 1200
DB 1141 CCAGCATCTTCAGAGCAGATGACCAAGATCTGTGAGCCCTTCCCGCGCCGCAACCCC 1200
QY 1201 GAGATCGTGATCTACCAAGGCCCTCTGTAGTGGGAGGACCTGGAGATCGGCGCAGC 1260
DB 1201 GAGATCGTGATCTACCAAGGCCCTCTGTAGTGGGAGGACCTGGAGATCGGCGCAGC 1260
QY 1261 CGGCGCAAGATCGAGGAGCTGCGCAAGCCTGCTGCCTGGGCTTCCACACCCCGCAG 1320
DB 1261 CGGCGCAAGATCGAGGAGCTGCGCAAGCCTGCTGCCTGGGCTTCCACACCCCGCAG 1320
QY 1321 AAGAACCCAGAGAGAGCCCTTCCCTGTGATGGGCTACGAGCTGACCCCGCAGCAAG 1380
DB 1321 AAGAACCCAGAGAGAGCCCTTCCCTGTGATGGGCTACGAGCTGACCCCGCAGCAAG 1380
QY 1381 TGGACCGTGAGCCCATCGAGCTGCCCGAGAGAGAGCTGGACCGTGAACGACATCCAG 1440
DB 1375 TGGACCGTGAGCCCATCGAGCTGCCCGAGAGAGAGCTGGACCGTGAACGACATCCAG 1434
QY 1441 AAGCTGTGGGCAAGCTGAATGGGCCAGCAGATCTACCCCGGCATCAAGGTGCGCCAG 1500
DB 1435 AAGCTGTGGGCAAGCTGAATGGGCCAGCAGATCTACCCCGGCATCAAGGTGCGCCAG 1494
QY 1501 CTGTGCAAGCTGTGCGCGGCGCAAGGCCCTGACCGACATCTGTCGCCCTGACCGAGGAG 1560
DB 1495 CTGTGCAAGCTGTGCGCGGCGCAAGGCCCTGACCGACATCTGTCGCCCTGACCGAGGAG 1554

QY 1561 GCCAGCTGTGAGCTGGCCGAGAAACCGGAGATCTCTGCGGAGCCCGTGCAGCGCTGTAC 1620
DB 1555 GCCAGCTGTGAGCTGGCCGAGAAACCGGAGATCTCTGCGGAGCCCGTGCAGCGCTGTAC 1614
QY 1621 TAGACCCCCAGCAAGACCTTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCTAGTGACC 1680
DB 1615 TAGACCCCCAGCAAGACCTTGGTGGCCGAGATCCAGAAGCAGGGCCACGACCTAGTGACC 1674
QY 1681 TACAGATCTTACCAGGAGCCCTTCAAGAACTGTAGACCCGGCAAGTACGCCAAGATCGGC 1740
DB 1675 TACAGATCTTACCAGGAGCCCTTCAAGAACTGTAGACCCGGCAAGTACGCCAAGATCGGC 1734
QY 1741 ACCGCCACACCAAGACGCTGAAGCAGCTACCGAGGCCGTGCAGAAGATCGCCATGGAG 1800
DB 1735 ACCGCCACACCAAGACGCTGAAGCAGCTACCGAGGCCGTGCAGAAGATCGCCATGGAG 1794
QY 1801 AGCATCTGTATCTGGGCAAGACCCCAAGTTCGCGCTGCCCATTCAGAAAGGACCTGG 1860
DB 1795 AGCATCTGTATCTGGGCAAGACCCCAAGTTCGCGCTGCCCATTCAGAAAGGACCTGG 1854
QY 1861 GAGACCTGGTGGACCGGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1920
DB 1855 GAGACCTGGTGGACCGGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGTTCGTGAAC 1914
QY 1921 ACCCCCCCTGTGTGAAGCTGTGTACACGCTGGAGAAGGAGCCCATCATCGGCGCCGAG 1980
DB 1915 ACCCCCCCTGTGTGAAGCTGTGTACACGCTGGAGAAGGAGCCCATCATCGGCGCCGAG 1974
QY 1981 ACCTTCTACGTGGACGGCGCCCAACCGCAGACCAAGATCGGCAAGGCGCGGTACGTG 2040
DB 1975 ACCTTCTACGTGGACGGCGCCCAACCGCAGACCAAGATCGGCAAGGCGCGGTACGTG 2034
QY 2041 ACCGACCGGGCGCGGCAAGATCTGTAGCCTGACCGAGACCCACCAAGGACCCGAG 2100
DB 2035 ACCGACCGGGCGCGGCAAGATCTGTAGCCTGACCGAGACCCACCAAGGACCCGAG 2094
QY 2101 CTGAGGCCATCCAGCTGGGCCCTGACGAGACAGCGGCGAGGTTGAACATCTGTGACCGAC 2160
DB 2095 CTGAGGCCATCCAGCTGGGCCCTGACGAGACAGCGGCGAGGTTGAACATCTGTGACCGAC 2154
QY 2161 AGCCAGTACGCCCTTGGGATCATCCAGGCCAGCCCGCAGAGCGAGGAGCGAGCTGGTG 2220
DB 2155 AGCCAGTACGCCCTTGGGATCATCCAGGCCAGCCCGCAGAGCGAGGAGCTGGTG 2214
QY 2221 AACCATGATCTGAGCAGCTGATCAAGAAAGAGAGTGTACCTGAGCTGGGTGCCGCC 2280
DB 2215 AACCATGATCTGAGCAGCTGATCAAGAAAGAGAGTGTACCTGAGCTGGGTGCCGCC 2274
QY 2281 CACAAGGGCATCGCGGCAACGAGCAGATCGACAAGCTGGTGGAGAGGGCATCCGCAAG 2340
DB 2275 CACAAGGGCATCGCGGCAACGAGCAGATCGACAAGCTGGTGGAGAGGGCATCCGCAAG 2334
QY 2341 GTGCTGTCTTGGACGCGCATCGATGGCGGCATCGTGATCTTACCAGTACATGGACGACTG 2400
DB 2335 GTGCTGTCTTGGACGCGCATCGATGGCGGCATCGTGATCTTACCAGTACATGGACGACTG 2394
QY 2401 TACGTGGGCGAGCGCGCCCTAGGATCGATTAAGGCTTCCCGGGGCTAGCACCGGTGAA 2460
DB 2395 TACGTGGGCGAGCGCGCCCTAGGATCGATTAAGGCTTCCCGGGGCTAGCACCGGTGAA 2454
QY 2461 TTC 2463
DB 2455 TTC 2457

RESULT 4
AX427930
LOCUS AX427930
DEFINITION Sequence 168 from Patent WO0232943.
ACCESSION AX427930
VERSION AX427930.1
KEYWORDS GI:21538017
SOURCE synthetic construct.

AX427930 9166 bp DNA linear PAT 20-JUN-2002

Best Local Similarity 91.3%; Pred. No. 2.8e-201;
Matches 2186; Conservative 0; Mismatches 187; Indels 22; Gaps 5;

Qy 14 TGGCGAGGCCATGAGCCAGG- --CCACCAGCGCAACATCTGATGCGAGCGAGCAACT 70
Db 2967 TGGCGAGGCCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGAGCGCGGCRAC 3026
Qy 71 TCAAGGGCCCAAGCGCATCATCAAGTCTCAACTGCGGCAAGAGGGCCACATCGGCC 130
Db 3027 TCCGCAACACGCGCAAGATCGTGAAGTCTCAACTGCGGCAAGAGGGCCACACCGCCC 3086
Qy 131 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTGGGGCAAGAGGGCCACACAGA 190
Db 3087 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTGGGGCAAGAGGGCCACACAGA 3146
Qy 191 TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGAGCTGGCCCTTCCCGCAGG 250
Db 3147 TGAAGGACTGCACCGAGCGGACAGGCTRA- TTTTATGGAAGATCTGGCCCTTCCCAAG 3205
Qy 251 GCAAGGCCGCGAGTTCCCGAGCGAGCAGAAACCGCGCCACAGCCCAACAGCGCGAGC 310
Db 3206 GGAAGGCCAGGGAATTTCTTCAGAGCAGACACAGAGCCACAGCCCAACAGAGAGAGC 3265
Qy 311 TGCAGGTGCGCG- ----CGACAAACCCCGCAGCGAGGCGCGCGAGCGCGAGGCA 364
Db 3266 TTCAGGTTTGGGGAAGAGACAACAACACTCCCTCTCAGAAAGCAGGAGCGGATAGACAAGAA 3325
Qy 365 -----CCCTGAACTTCCCGCAGATCACTTGTGGCAGCGCCCTGCTGAGCATCAAG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCAGCCCTCGTCAATAAGA 3395
Qy 419 TGGCGGCGCAGATCAAGAGGCGCCTGTGTGACACCGCGCGGAGCAGACCGTGTGAGG 478
Db 3386 TAGGGGCGCAGCTGAAGAGGCGCCTTCTAGACACCGCGCGGAGCAGACCGTGTGAGG 3445
Qy 479 AGATGACCTTGCCCGCAAGTGGNAGCCCAAGATGATCGGGGGATCGGGGCTTATCA 538
Db 3446 AGATGAACTTGCGCGCGCGTGGAAAGCCCAAGATGATCGGGGCGATCGGGGCTTCA 3505
Qy 539 AGGTGCGCCAGTAGCAGACAGATCTGTGATCGAGATCTCGGCAAGAAAGCCATCGGCACCG 598
Db 3506 AGGTGCGCCAGTAGCAGACAGATCTGTATCGAGATCTCGGCGCACAGGCCATCGGCACCG 3565
Qy 599 TGTGATGCGCGCCACCGCCGTGAACATCATCGGCGCAACATGCTGACCCAGCTGGCT 658
Db 3566 TGTGTGTGGCGCCACCGCCGTGAACATCATCGGCGCAACCTGCTGACCCAGATCGCT 3625
Qy 659 GCACCTGAACTTCCCATCAGCCCATCAGACCGTGAAGCGTGGCGGTGAAGCTGAAGCCCGCA 718
Db 3626 GCACCTGAACTTCCCATCAGCCCATCAGACCGTGAAGCGTGGCGGTGAAGCTGAAGCCCGCA 3685
Qy 719 TGAAGGCCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCG 778
Db 3686 TGAAGGCCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGTTGG 3745
Qy 779 CCATCTCGAGAGGATGGAAGAGGAGGCAAGATCAACAGATCGGCGCCCGAGAACCCCT 838
Db 3746 AGATCTGCACCGAGATGGAAGAGGAGGCAAGATCAGCAAGATCGGCGCCCGAGAACCCCT 3805
Qy 839 ACAACACCCCGTGTGCGCATCAAGAAGAGGACACCAAGTGGCGCAAGCTGGTGG 898
Db 3806 ACACACCCCGTGTGCGCATCAAGAAGAGGAGGCAAGATCAGCAAGTGGCGCGCAAGCTGGTGG 3865
Qy 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAAGTGGGATCCCC 958
Db 3866 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAAGTGGGATCCCC 3925
Qy 959 ACCCGCGCGCTGAAGAAGAAGAGCGCTGACCGTGTGAGCGTGGGCGAGCGCTACT 1018
Db 3926 ACCCGCGCGCTGAAGAAGAAGAGCGCTGACCGTGTGAGCGTGGGCGAGCGCTACT 3985
Qy 1019 TCAGCGTGGCCCTGGAGGAGCTTCGCAAGTACACCGCCTTCAACATCCCCAGCATCA 1078
Db 1019 TCAGCGTGGCCCTGGAGGAGCTTCGCAAGTACACCGCCTTCAACATCCCCAGCATCA 1078

Db 3986 TCAGCGTGGCCCTGGACAAGGACTTCCGCAAGTACACCGCCCTTCAACATCCCCAGCATCA 4045
Qy 1079 ACAAGAGACCCCGGCATCCGCTACCAAGTACAAGTCTGCGCCAGGCTGGAGGGCA 1138
Db 4046 ACAAGAGACCCCGGCATCCGCTACCAAGTACAAGTCTGCGCCAGGCTGGAGGGCA 4105
Qy 1139 GCGCCAGCATTTCCAGAGCAGCATGACCAAGATCCTTGAGAGCCCTTCCGCGCCCGCAACC 1198
Db 4106 GCGCCGCACTTTCCAGTGCAGCATGACCAAGATCCTTGAGAGCCCTTCCGCAAGCAGAACC 4165
Qy 1199 CCGAGATCGTATCTACCA-----GGCCCCCTGTAGTGGGAGCAGCCTGGAGATCG 1252
Db 4166 CCGACATCGTATCTACCAAGTACATGAGACCACTGTAGTGGGAGCAGCCTGGAGATCG 4225
Qy 1253 GCGAGCAGCGCGCAAGATCGAGGAGTGCACAGCAGCTGCTGCGTGGGCTTCACCA 1312
Db 4226 GCGAGCAGCGCACCAAGATCGAGGAGTGCAGGAGCTGCGGAGCAGCCTGCTGCGTGGGCTTCACCA 4285
Qy 1313 CCGCCGACAGAAGCACCAGAAGGAGCCCTTCTTCTGTGTGATGGGTACGAGTGCACC 1372
Db 4286 CCGCCGACAGAAGCACCAGAAGGAGCCCTTCTTCTGTGTGATGGGTACGAGTGCACC 4345
Qy 1373 CCGACAAGTGCAGCTGCAGCCCATGAGCTGCGCGGAGAGAGAGTGGACCGTGAACG 1432
Db 4346 CCGACAAGTGCAGCTGCAGCCCATGAGCTGCGCGGAGAGAGTGGACCGTGAACG 4405
Qy 1433 ACATCCAGAAGCTGTTGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGCATCAAG 1492
Db 4406 ACATCCAGAAGCTGTTGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGCATCAAG 4465
Qy 1493 TGCGCCAGCTGTGCAAGCTGCTGCGCGCGCCCAAGGCGCTTACCGAGATCTGTCGCCCTGA 1552
Db 4466 TGCGCCAGCTGTGCAAGCTGCTGCGCGCACCAGGCGCTTACCGAGTGTGTCGCCCTGA 4525
Qy 1553 CCGAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTCGGCGAGCCCGTGCACG 1612
Db 4526 CCGAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTCGGCGAGCCCGTGCACG 4585
Qy 1613 GCGTGTACTAGGACCCAGCAAGGAGCTGTTGGCGGAGATCCAGAACGAGGCGCACGACC 1672
Db 4586 GCGTGTACTAGGACCCAGCAAGGAGCTGTTGGCGGAGATCCAGAACGAGGCGCACGACC 4645
Qy 1673 AGTGGACCTACAGATCTACCGAGGAGCTTCAAGAACCTTGAAGACCGGCGCAAGTACGCA 1732
Db 4646 AGTGGACCTACAGATCTACCGAGGAGCTTCAAGAACCTTGAAGACCGGCGCAAGTACGCA 4705
Qy 1733 AGATGCGCACCGGCCACCAACGACGTGAAGAGCTGACCGAGGCGCTGCAGAGATCG 1792
Db 4706 GCATGAAGGCGGCCACCAACGACGTGAAGAGCTGACCGAGGCGCTGCAGAGATCG 4765
Qy 1793 CCATGGAGAGCATCGTATCTGGGCGAAGACCCCAAGTTCGCCCTGCCATCCAGAGG 1852
Db 4766 CCACCGAGAGCATCGTATCTGGGCGAAGACCCCAAGTTCAGCTGCCATCCAGAGG 4825
Qy 1853 AGACCTGGGAGACCTGGTGGACCGACTACTTGGCAGGCGACCTTGATCCCGAGTGGAGT 1912
Db 4826 AGACCTGGGAGCGCTGGTGGACCGAGTACTTGGCAGGCGACCTTGTATCCCGAGTGGAGT 4885
Qy 1913 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTATCAGCTGGAAGAGAGGCGCCATCATCG 1972
Db 4886 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTATCAGCTGGAAGAGAGGCGCCATCATCG 4945
Qy 1973 GCGCCGAGAGCTTCTAGTGGAGCGCGCCCAACCGCAGACCAAGATCGGCAAGGCGC 2032
Db 4946 GCGCCGAGAGCTTCTAGTGGAGCGCGCCCGCCCAACCGCAGACCAAGCTGGCAAGGCGC 5005
Qy 2033 GCTACGCTGACCGCGCGCGCGCGCAAGATCGTGAAGCTTACCGAGACCAACCAACGAG 2092
Db 5006 GCTACGCTGACCGCGCGCGCGCGCAAGATCGTGAAGCTTACCGAGACCAACCAACGAG 5065
Qy 2093 AGACCGAGCTGCGAGGCGCATCCAGCTGCCCTTGCAGGACAGCGGCGAGGCTGAACATCG 2152
Db 5066 AGACCGAGCTGCGAGGCGCATCCAGCTGCCCTTGCAGGACAGCGGCGCTGGAGGTGAACATCG 5125

QY 2153 TGACGACAGCCAGTAGCCCTGGGCTATCATCCAGGCCAGCCGACAAAGAGGAGAGCG 2212
Db 5126 TGACGACAGCCAGTAGCCCTGGGCTATCATCCAGGCCAGCCGACAAAGAGGAGAGCG 5185
QY 2213 AGCTGGTGAACCAAGATCATCGACAGCTGTATCAAGAGGAGAGAGGTGTACTGAGCTGGG 2272
Db 5186 AGCTGGTGAACCAAGATCATCGACAGCTGTATCAAGAGGAGAGAGGTGTACTGAGCTGGG 5245
QY 2273 TGCCGCCACACAGGAGATCGGGGGCAACGAGAGAGATCGACAGCTGTGTAGCAAGGGCA 2332
Db 5246 TGCCGCCACACAGGAGATCGGGGGCAACGAGAGAGATCGACAGCTGTGTAGCAAGGGCA 5305
QY 2333 TCCGCAAGGTGCTGTCTTCCGAGGAGATCGATGGCGGATCGTGTATACCACTA 2387
Db 5306 TCCGCAAGGTGCTGTCTTCCGAGGAGATCGATGGCGGATCGATGGCGGAGGAGGAGGAGGAG 5360

RESULT 8
AX427925
LOCUS AX427925 9194 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 163 from Patent WO0232943.
ACCESSION AX427925
VERSION AX427925.1 GI:21538012
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequences.
1
REFERENCE
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 163 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source Location/Qualifiers
1..9194
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"

BASE COUNT 2232 a 2671 c 2524 g 1767 t
ORIGIN

Query Match 81.2%; Score 2000.2; DB 6; Length 9194;
Best Local Similarity 91.2%; Pred. No. 4.1e-201;
Matches 2185; Conservative 0; Mismatches 188; Indels 22; Gaps 5;

QY 14 TGGCGAGGCGCATGAGCCAGG---CCACAGCGCGCAACATCTGTATCGACGGCAGCAACT 70
Db 2967 TGGCGAGGCGCATGAGCCAGGTCACCAACAGCGCGCAACATCTGTATCGACGGCAGCAACT 3026
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGTCTCAACTCGCGGAGGAGGCGCACATCGCCC 130
Db 3027 TCCGCAACAGCGCAAGATCGTGAAGTGTCTCAACTCGCGGAGGAGGCGCACATCGCCC 3086
QY 131 GCAACTCGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGGAGGAGGCGCCACAGGA 190
Db 3087 GCAACTCGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGGAGGAGGCGCCACAGGA 3146
QY 191 TGAAGGACTGACGAGCGCGCCAGGCGCAACTCTTCCGCGAGGAGCTGGCTTCCCCCAGG 250
Db 3147 TGAAGGACTGACGAGCGAGGAGGCTAA-TTTTATTAGGGAAGATCTGGCCCTTCCCAAG 3205
QY 251 GCAAGGCGCGGAGTCTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 310
Db 3206 GGAAGGCGAGGAGATTTCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3255
QY 311 TGCAGGTCGCGGG-----CGACAACCCCGCGAGCGAGGCGCGGAGGCGGAGGCGCA 364
Db 3266 TTCAGGTTTGGGGAAGAGAGCAACAACATCCCTCTCAGAAGGAGGAGGAGGAGGAGGAGG 3325
QY 365 -----CCCTGAACCTTCCCGAGATCAACCTGTGGGAGGCGGCGGCTGGTGAGCATCAAG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTGTCACATAAAGA 3385

QY 419 TGGCGGCGCAGATCAAGGAGGCGCTGCTGCACACCGCGCGCGGACGACACCGCTGCTGGAGG 478
Db 3386 TAGGGGCGCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGGAGGAGACCGCTGCTGGAGG 3445
QY 479 AGATGAGCCTTCCCGGCAAGTGGAAAGCCCAAGATGATCGCGGCGATCGCGGCTTTCATCA 538
Db 3446 AGATGAACCTTCCCGGCGCTGGAAAGCCCAAGATGATCGCGGCGATCGCGGCTTTCATCA 3505
QY 539 AGTGCGCGCAGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 598
Db 3506 AGTGCGCGCAGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3565
QY 599 TGTGTGATCGGCGCCACCGCTGAAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT 658
Db 3566 TGTGTGATCGGCGCCACCGCTGAAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT 3625
QY 659 GCACCTGAACTTCCCGCATCAGCGCCCATCGAGAGCGCTGCGCGTGAAGCTGAAGCGCGGCA 718
Db 3626 GCACCTGAACTTCCCGCATCAGCGCCCATCGAGAGCGCTGCGCGTGAAGCTGAAGCGCGGCA 3685
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCGCCCTGACCGGAGGAGAGATCAAGCGCCCTGACCG 778
Db 3686 TGGACGCGCCCAAGGTGAAGCAGTGGCGCCCTGACCGGAGGAGAGATCAAGCGCCCTGACCG 3745
QY 779 CCATCTCGGAGGAGATGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 838
Db 3746 AGATCTGCACCGAGATGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3805
QY 839 ACAACACCCCGCTTTCGCCCATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 898
Db 3806 ACAACACCCCGCTTTCGCCCATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3865
QY 899 ACTTCCGCGAGCTGAACAAGCGGACCCAGGAGCTTCTGGGAGGTCAGCTGGGCGATCCCC 958
Db 3866 ACTTCCGCGAGCTGAACAAGCGGACCCAGGAGCTTCTGGGAGGTCAGCTGGGCGATCCCC 3925
QY 959 ACCCGCGCGCTGAAAGAAAGAGAGCGCTGACCGTGTGAGCTGGGCGGCGACCGCTACT 1018
Db 3926 ACCCGCGCGCTGAAAGAAAGAGAGCGCTGACCGTGTGAGCTGGGCGGCGACCGCTACT 3985
QY 1019 TCAGCTGCGCCCTGGAGGAGGACTTCCGCAAGTACACCGGCTTCCACATCCCCAGCATCA 1078
Db 3986 TCAGCTGCGCCCTGGAGAGGAGCTTCCGCAAGTACACCGGCTTCCACATCCCCAGCATCA 4045
QY 1079 ACACGAGAGCCCCCGGATCCGCTACCAAGTACAAAGCTGCTGCCCGAGGCTGGAGGAGCA 1138
Db 4046 ACACGAGAGCCCCCGGATCCGCTACCAAGTACAAAGCTGCTGCCCGAGGCTGGAGGAGCA 4105
QY 1139 GCGCCAGCATCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1198
Db 4106 GCGCCAGCATCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4165
QY 1199 CCGAGATCGTGAATACCA-----GGCCCCCTGTAGTGGGAGGAGGAGGAGGAGGAGGAGG 1252
Db 4166 CCGAGATCGTGAATACCA-----GGCCCCCTGTAGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 4225
QY 1253 GCCAGACCGCGCCCAAGATCGAGGAGTGGCGAAGCAGCTGCTGGCTGGGCTTTCACCA 1312
Db 4226 GCCAGACCGCGCCCAAGATCGAGGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4285
QY 1313 CCGCGGACAGAAAGCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1372
Db 4286 CCGCGGACAGAAAGCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4345
QY 1373 CCGACAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1432
Db 4346 CCGACAGTGGACCGTGCAGCCCATCGCTGCTGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4405
QY 1433 ACATCCAGAGGCTGGTGGGCAAGCTGAACCTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1492
Db 4406 ACATCCAGAGGCTGGTGGGCAAGCTGAACCTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4465

QY 1493 TGGCCAGCTGTGCAAGCTGTGCGCGCGCCCAAGGCGCTGACCGACATCGTGCCTCGTA 1552
Db TGGCCAGCTGTGCAAGCTGTGCGCGCGCCCAAGGCGCTGACCGACATCGTGCCTCGTA 4525
QY 1553 CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCCCGTGCACG 1612
Db CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCCCGTGCACG 4585
QY 1613 GCGTGTACTAGGACCCAGCAAGGACCTGTGCGCGAGATCCAGAACGAGGCGCACACC 1672
Db GCGTGTACTAGGACCCAGCAAGGACCTGTGCGCGAGATCCAGAACGAGGCGCACAGGCC 4645
QY 1673 AGTGGACCTACCAAGCTTACCAAGAGCGCTTCAAGAACCTGAAGACCGCAAGTACGCCA 1732
Db AGTGGACCTACCAAGCTTACCAAGAGCGCTTCAAGAACCTGAAGACCGCGCAAGTACGCC 4705
QY 1733 AGATGCCACCGCCACCAACAGACGTGAAGAGCTGACCGAGGCGGTGCAAGAATCG 1792
Db AGATGCCACCGCCACCAACAGACGTGAAGAGCTGACCGAGGCGGTGCAAGAATCG 4765
QY 1793 CCATGGAGAGCATCGTGATCTGGGGAGAACCCCGAAGTTCCCGCTGCCCATCCAGAGG 1852
Db CCATGGAGAGCATCGTGATCTGGGGAGAACCCCGAAGTTCCCGCTGCCCATCCAGAGG 4825
QY 1853 AGACCTGGGAGACCTGTGGACCGACTACTGGCAGGCGCACTGGATCCCGAGTGGGAGT 1912
Db AGACCTGGGAGACCTGTGGACCGAGTACTGGCAGGCGCACTGGATCCCGAGTGGGAGT 4885
QY 1913 TCGTGAACACCCCGCTGTGTAAGCTGTGTAACAGCTGGAGAGAGGCCCATCATCG 1972
Db TCGTGAACACCCCGCTGTGTAAGCTGTGTAACAGCTGGAGAGAGGCCCATCATCG 4945
QY 1973 GCGCGGAGACCTTCTAGCTGGAGCGCGCCCAAGCCGAGACCAAGATCGCAAGGCCG 2032
Db GCGCGGAGACCTTCTAGCTGGAGCGCGCCCAAGCCGAGACCAAGATCGCAAGGCCG 5005
QY 2033 GCTAGCTGACCGAGCGCGCGCGCGCAGAGATCGTGAAGCTGACCGAGACCAACACCA 2092
Db GCTAGCTGACCGAGCGCGCGCGCGCAGAGATCGTGAAGCTGACCGAGACCAACACCA 5065
QY 2093 AGACCGAGCTGACGAGCATCCAGCTGGCGCTGAGGAGCGGCGCGCTGAGGAGTGAACATCG 2152
Db AGACCGAGCTGACGAGCATCCAGCTGGCGCTGAGGAGCGGCGCGCTGAGGAGTGAACATCG 5125
QY 2153 TGACCGACAGCCAGTACGCGCTGGGACATCCAGGCGCCAGCGCGCGCGAGAGCG 2212
Db TGACCGACAGCCAGTACGCGCTGGGACATCCAGGCGCCAGCGCGCGAGAGCG 5185
QY 2213 AGCTGTGAACAGATCATCGACGCTGATCAAGAGGAGAGAGTGTACCTGAGCTGGG 2272
Db AGCTGTGAACAGATCATCGACGCTGATCAAGAGGAGAGAGTGTACCTGAGCTGGG 5245
QY 2273 TGCGCGCCCAACAGGCGATCGCGCGCAACGAGCAGATCGACAGCTGGTGAAGGGCA 2332
Db TGCGCGCCCAACAGGCGATCGCGCGCAACGAGCAGATCGACAGCTGGTGAAGGGCA 5305
QY 2333 TCCGCAAGGTGCTTCTCCTGGAGCGCATCGATGGCGGCGCATCGATCGTACACGTA 2387
Db TCCGCAAGGTGCTTCTCCTGGAGCGCATCGATGGCGGCGCATCGATCGTACACGTA 5360

RESULT 9
AX427927
LOCUS AX427927 12411 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 165 from Patent W00232943.
ACCESSION AX427927
VERSION AX427927.1 GI:21538014
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE artificial sequences.
1
AUTHORS Huang, Y. and Nabel, G.J.

TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 165 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES Location/Qualifiers
source 1..12411
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 2938 a 3733 c 3476 g 2264 t
ORIGIN
Query Match 81.2%; Score 2000.2; DB 6; Length 12411;
Best Local Similarity 91.2%; Pred. No. 3.8e-201;
Matches 2185; Conservative 0; Mismatches 188; Indels 22; Gaps 5;
QY 14 TGGCCGAGGCGCATGAGCCAGG---CCACCAGCGCAACATCTCTGATGACGCGCAGCAACT 70
Db TGGCCGAGGCGCATGAGCCAGGCGCATGAGCCAGGCGCAACATCTCTGATGACGCGCAGCAACT 3026
QY 71 TCAAGGCGCCCAAGCGGATCATCAAGTGTCTCAACTCGGCAAGAGGGCCACATCGGCC 130
Db TCAAGGCGCCCAAGCGGATCATCAAGTGTCTCAACTCGGCAAGAGGGCCACATCGGCC 3086
QY 131 GCAACTGCCCGCGCCCGCCCAAGAGGGCTGTGGAAGTGTCTCAACTCGGCAAGAGGGCCACAC 190
Db GCAACTGCCCGCGCCCGCCCAAGAGGGCTGTGGAAGTGTCTCAACTCGGCAAGAGGGCCACAC 3146
QY 191 TGAAGGACTCACCGAGCGCCAGCCAACTTCTTCCCGAGGAGCTTGGCCCTTCCCGCAGG 250
Db TGAAGGACTCACCGAGCGACAGGCTAA-TTTTATTAGGAAGATCTGGCTTCCCAACAG 3205
QY 251 GCAAGGCGCCCGAGTTCGCCAGCGAGCAGAGAACCGCGCCACACGCCCCACAGCGCGCAGC 310
Db GCAAGGCGCCCGAGTTCGCCAGCGAGCAGAGAACCGCGCCACACGCCCCACAGAGAGAGC 3265
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCA 364
Db TGCAGGTGCGCGGAGAGAGAGCAACTCCCTCTCAGAGCAGAGCGCGCGCGCGCGCGCA 3325
QY 365 -----CCCTGAATCTCCCGAGATCACTCTGTGGCAGCGCGCGCGCGCGCGCGCG 418
Db -----CCCTGAATCTCCCGAGATCACTCTGTGGCAGCGCGCGCGCGCGCGCGCGCG 3385
QY 419 TGGCGCGCAGATCAAGAGGCGCTGTGTGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 478
Db TGGCGCGCAGATCAAGAGGCGCTGTGTGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 3445
QY 479 AGATGAGCCTGCCCGGCAAGTGAAGCCCAAGATGATCGCGCGCGCGCGCGCGCGCGCG 538
Db AGATGAGCCTGCCCGGCG 3505
QY 539 AGTGGCGCAGTACGACCATCTCTGATCGAGATCTGTGGCAAGAGAGGCGCATCGGCACCG 598
Db AGTGGCGCAGTACGACCATCTCTGATCGAGATCTGTGGCAAGAGAGGCGCATCGGCACCG 3565
QY 599 TGTGTGATCG 658
Db TGTGTGATCG 3625
QY 659 GCACCTGAACCTTCCCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 718
Db GCACCTGAACCTTCCCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3685
QY 719 TGGACG 778
Db TGGACG 3745
QY 779 CCATCTGCGAGGAGATGGAAGAGGCGGCAAGATCCACCAAGATCGCGCGCGCGCGCG 838
Db CCATCTGCGAGGAGATGGAAGAGGCGGCAAGATCCACCAAGATCGCGCGCGCGCGCG 3805
QY 839 ACAACACCG 898

Db 3806 ACAACACCCCGTGTTCGCCATCAAGAAGACAGCACCAGTGGCGCAAGCTGGTG 3865
QY 899 ACTTCCGCGAGCTGAACAAAGCAGCCAGGACTTCTGGGAGGTGAGCTGGGATCCCCC 958
Db 3866 ACTTCCGCGAGCTGAACAAAGCAGCCAGGACTTCTGGGAGGTGAGCTGGGATCCCCC 3925
QY 959 ACCCGCGCGCTGAAGAAGAAGAGCGGTGACCCGTGCTGGAGCTGGCGAGCGCTACT 1018
Db 3926 ACCCGCGCGCTGAAGCAGAAGAGAGCGGTGACCCGTGCTGGAGCTGGCGAGCGCTACT 3985
QY 1019 TCAGGTGCCCCGTGACGAGGAGTTCGGAAGTACACCGCTTACCATCCCAAGCATCA 1078
Db 3986 TCAGGTGCCCCGTGACGAGGAGTTCGGAAGTACACCGCTTACCATCCCAAGCATCA 4045
QY 1079 ACAAGAGACCCCGCGCATCCCGTACCACTACAAGTGTGCTGCCAGGCTGGAAGGCA 1138
Db 4046 ACAAGAGACCCCGCGCATCCCGTACCACTACAAGTGTGCTGCCAGGCTGGAAGGCA 4105
QY 1139 GCCCAGACNTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACC 1198
Db 4106 GCCCGCCCATCTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGAAC 4165
QY 1199 CCGAGATCGTGATCTACCA- ----GGCCCCCTGTACGTGGGAGCGACCTGGAGATCG 1252
Db 4166 CCGATCGTGATCTACCAAGTACATGGAGCGACCTGTACGTGGGAGCGACCTGGAGATCG 4225
QY 1253 GCCAGACCGCCCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGGCTTTCACCA 1312
Db 4226 GCCAGACCGCCCAAGATCGAGGAGCTGCGCAGACCTGCTGGGCTTTCACCA 4285
QY 1313 CCCCCAGAACGACCAAGAGAGGAGGCCCCCTTCTGTGGATGGGCTACGAGCTGCACC 1372
Db 4286 CCCCCAGAACGACCAAGAGAGGAGGCCCCCTTCTGTGGATGGGCTACGAGCTGCACC 4345
QY 1373 CCGACAAGTGAACGCTGAGCCATCGAGCTGCCCCAGAGAGAGCTGGACCGTGAAG 1432
Db 4346 CCGACAAGTGAACGCTGAGCCATCGTGTGCTGCCGAGAGGAGAGCTGGACCGTGAAG 4405
QY 1433 ACATCCAGAAGCTGTGGGCAAGCTGAATGGGCGCAGCAGATCTACCCCGCATCAAGG 1492
Db 4406 ACATCCAGAAGCTGTGGGCAAGCTGAATGGGCGCAGCAGATCTACCCCGCATCAAGG 4465
QY 1493 TCGCCACCTGTGAAGCTGTGCGCGCGGCGAGCCCTGACGACATCGTGCCCTGA 1552
Db 4466 TCGCGCAGCTGTGAAGCTGTGCGCGCAGCAGGCGCTGACGAGTGTGTCGCCCTGA 4525
QY 1553 CCGAGGAGCGGAGCTGGAGCTGGCGGAGACCGGAGATCTGCGGAGCCCGTGCAGG 1612
Db 4526 CCGAGGAGCGGAGCTGGAGCTGGCGGAGACCGGAGATCTGAGGAGCCCGTGCAGG 4585
QY 1613 GCGTGTACTAGACCCCAAGCAAGCACTGTGTGGCGGAGATCCAGAGCAGGCGCCAGCACC 1672
Db 4586 GCGTGTACTAGACCCCAAGCAAGCACTGTGTGGCGGAGATCCAGAGCAGGCGCCAGG 4645
QY 1673 AGTGGACCTACAGATCTACAGAGGAGCCCTTCAAGACCTGAAGACCGGCGAGTACGCCA 1732
Db 4646 AGTGGACCTACAGATCTACAGAGGAGCCCTTCAAGACCTGAAGACCGGCGAGTACGCCA 4705
QY 1733 AGATGCGCAGCCCAACACCAAGCAGTGAAGCAGCTGACCGAGCGCGTGCAGAGATCG 1792
Db 4706 GCATGAAGGGGCGCCACACCAAGCAGTGAAGCAGCTGACCGAGCGCGTGCAGAGATCG 4765
QY 1793 CCATGGAGAGATCGTGATCTGGGCGAAGACCCCAAGTTCGCGCTGCCATCCAGAGG 1852
Db 4766 CCACGAGAGATCGTGATCTGGGCGAAGACCCCAAGTTCGCGCTGCCATCCAGAGG 4825
QY 1853 AGACCTGGAGACCTGCTGGACCACTACTTGGGAGGCGACCTGTGATCCCGAGTGGAGT 1912
Db 4826 AGACCTGGGAGGCGCTGCTGGACCAAGTACTTGGCAGGCGACCTGTGATCCCGAGTGGAGT 4885
QY 1913 TCGTGAACACCCCGCTGCTGAAGCTGGTACCAAGCTGGAGAGGAGGCGCCATCATCG 1972

Db 4886 TCGTGAACACCCCGCTGCTGAAGCTGTGTACCAAGTGGAGAGGAGCCCATCATCG 4945
QY 1973 GCGCCGAGACCTTCTACGTGGACGGCCGCCCAACCGCGAGACCAAGATCGGCAAGGCCG 2032
Db 4946 GCGCCGAGACCTTCTACGTGGACGGCCGCCCAACCGCGAGACCAAGTGGCAAGGCCG 5005
QY 2033 GCTACGTGACCGACCGGCGCGCAGAGATCTGAGCCCTGACCGAGACCAACAGCAGA 2092
Db 5006 GCTACGTGACCGACCGGCGCGCAGAGGTGTGTGCTGACCGACCAACAGCAGA 5065
QY 2093 AGACCGAGCTGACGAGCCATCCAGCTGGCCCTGCAGGACAGCGAGCTGAACATCG 2152
Db 5066 AGACCGAGCTGACGAGCCATCCAGCTGGCCCTGCAGGACAGCGGCTGGAGTGAACATCG 5125
QY 2153 TGACCGACAGCAGTACGAGCCCTGGGCGCATCATCAGGCGCCAGCCGACAAAGAGCAGCG 2212
Db 5126 TGACCGACAGCAGTACGAGCCCTGGGCGCATCATCAGGCGCCAGCCGACAAAGAGCAGCG 5185
QY 2213 AGCTGTGAACAGATCATCGAGCAGCTGATCAAGAGAGAGAGTGTACCTGAGCTGG 2272
Db 5186 AGCTGTGAACAGATCATCGAGCAGCTGATCAAGAGAGAGAGTGTACCTGAGCTGG 5245
QY 2273 TGCCCGCCCAAGGCGCATCGGCGCAGCAGATCGAGCCCGAGCCGAGCAAGGGCA 2332
Db 5246 TGCCCGCCCAAGGCGCATCGGCGCAGCAGCAGCTGAGCGGCTGGTGGCGCGGCA 5305
QY 2333 TCCGCAAGTGTCTTCTGGAGCGCATCGAGCGCGCATCGTGATCTACCAAGTA 2387
Db 5306 TCCGCAAGTGTCTTCTGGAGCGCATCGAGCGCGCATCGAGCGCGCGAGGAGCAGGAA 5360

RESULT 10
AX427938
LOCUS AX427938 9785 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 176 from Patent WO0232943.
ACCESSION AX427938
VERSION AX427938.1 GI:21538025
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 176 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source
1. 9785
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2360 a 2843 c 2731 g 1851 t
ORIGIN

Query Match 81.1%; Score 1996.8; DB 6; Length 9785;
Best Local Similarity 91.3%; Pred. No. 9.1e-201;
Matches 2183; Conservative 0; Mismatches 182; Indels 27; Gaps 5;

QY 14 TGGCGAGGCGCATGAGCCAGGCGCACCAACATCTGATGCGGCGCAGCAACTTCA 73
Db 2967 TGGCGAGGCGCATGAGCCAGGTGCGAGCACACCAACATCATGATGAGCGGCACTTCA 3026
QY 74 AGGCGCCCAAGCGCATCATCAAGTGTCTCACTGCGGCAAGAGGCGCACATCGCCCCGA 133
Db 3027 AGGCGCCCAAGCG---CATCAAGTGTCTCACTGCGGCAAGAGGCGCACCTTGGCCGCA 3083
QY 134 ACTGCGCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGGCAAGAGGCGCACACCATGGA 193
Db 3084 ACTGCGCGCGCTCGGCAAGAGGCTGCTGGAAGTGGGCAAGAGGCGCACACCATGGA 3143
QY 194 AGGACTGACCGAGCGCGCAGGCCAACTTCTTCGCGAGGAGCTGGCCCTTCCCGCAGGCA 253

Db 3144 AGGACTGCACCGGAGGACAGGCTAA-----TAGGGAAAGATCTGGCCCTTCCCAACAGGGA 3197
Qy 254 AGGCCCCGAGTTCCTCCAGGAGAGCAAGCCGCCAACAGCCCAACACACAGCCGCGAGGTGC 313
Db 3198 AGGCCAGGGAATTCTTCAGAGCAGACAGACAGCAACAGCCCAACAGAGAGAGCTTC 3257
Qy 314 AGTGTCCCGG-----CGACAACCCCGCAGAGAGGCCGCGGCGAGCGCCAGGCA--- 364
Db 3258 AGGTTTGGGGAAGAGACAACAATCTCTCAGAAAGAGAGCGGATAGACAAGGAATG 3317
Qy 365 ---CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTTGGTGAGCATCAAGGTG 421
Db 3318 TATCTTTAGTTCCTTCAGATCACTTTTGGCAGCAGCCCTCGTCACATAAAGATAG 3377
Qy 422 CGGCCAGATCAAGGAGGCCCTGCTGACACCGGCGCCGACGACACCGTGTGGAGGAGA 481
Db 3378 GGGCCAGCTGAAGGAGGCCCTTCTAGACACCGGCGCGACGACACCGTGTGGAGAGA 3437
Qy 482 TGACCTTGGCCGCAAGTGGAAAGCCCAAGATGATCGGGGCATCGGGGCTTCATCAAG 541
Db 3438 TGAACCTTGGCCGCGGCGCTGGAAGCCCAAGATGATCGGGGCGATCGGGGCTTCATCAAG 3497
Qy 542 TGCGCCAGTACGACAGATCTGATCAGATCTCGGCAAGAGGCCATCGGCACCTGC 601
Db 3498 TGGCCAGTACGACAGATCTGATCGAGATCTCGGGCCACAAGGCCATCGGCACCTGC 3557
Qy 602 TGATCGGCCCAACCGGCGTGAACATCATCGGGCCGCAACATGCTGACCCAGTGGGTGCA 661
Db 3558 TGGTGGGCCCAACCGGCGTGAACATCATCGGGCCGCAACCTGCTGACCCAGATCGGCTGCA 3617
Qy 662 CCCTGAACTTCCCATCAGCCCACTGAGACCGTGGCCGCTGAAGTGAAGCCCGGCATGG 721
Db 3618 CCCTGAACTTCCCATCAGCCCACTGAGACCGTGGCCGCTGAAGTGAAGCCCGGCATGG 3677
Qy 722 ACGGCCCAAGTGAAGCAGTGGCCCTGTACCGAGGAGAGATCAAGGCCCTGACCGCA 781
Db 3678 ACGGCCCAAGTGAAGCAGTGGCCCTGTACCGAGGAGAGATCAAGGCCCTTGGTGA 3737
Qy 782 TCTGGAGGAGATGGAGAGAGGCGACAGATCACAGATCGGCCCGCGAGAACCCCTACA 841
Db 3738 TCTGACCGAGATGGAGAGAGGCGCAAGATCAGCAGATCGGCCCGCGAGAACCCCTACA 3797
Qy 842 ACACCCCGCTTTCGCCATCAAGAAAGAGGACAGCACCAGTGGCGCAAGCTGGTGACT 901
Db 3798 ACACCCCGCTTTCGCCATCAAGAAAGAGGACAGCACCAGTGGCGCAAGCTGGTGACT 3857
Qy 902 TCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGACGTGGGCGATCCCCAC 961
Db 3858 TCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGACGTGGGCGATCCCCAC 3917
Qy 962 CCGCGGCTTGAAGAAAGAGAGCGTGACCGTGTGGAGTGGCGGACCGCTACTTCA 1021
Db 3918 CCGCGGCTTGAAGAAAGAGAGCGTGACCGTGTGGAGTGGCGGACCGCTACTTCA 3977
Qy 1022 GCGTGCCTTGGACGAGGACTTCGCCAAGTACACCGCTTCCACCATCCCCAGCATCAACA 1081
Db 3978 GCGTGCCTTGGACAAAGACTTCGCCAAGTACACCGCTTCCACCATCCCCAGCATCAACA 4037
Qy 1082 ACGAGACCCCGGCGATCCGCTACAGTACAGTGTCTGCCCGAGGGCTGGAGGGCAGCC 1141
Db 4038 ACGAGACCCCGGCGATCCGCTACAGTACAGTGTCTGCCCGAGGGCTGGAGGGCAGCC 4097
Qy 1142 CCAGCATCTTCAGAGCAGATGACCAAGATCTCTGGAGCCCTTCGCGCCCGCAACCCCG 1201
Db 4098 CCGCCATCTTCAGTGTACAGTACAGTGTCTGGAGCCCTTCGCGAAGCAGAACCCCG 4157
Qy 1202 AGATCGTGTATTCACCA-----GGCCCCCTGTAGTGGGCGAGCAGCTGGAGATCGGCC 1255
Db 4158 ACATCGTGTATTCACAGTACATGAGCACCCTGTACGTGGGCGAGCAGCTGGAGATCGGCC 4217
Qy 1256 AGCACCAGCCAGATTCAGAGAGTGGCGAAGCAGCTGCTGGCTGGGGCTTCACCAACC 1315
Db 4218 AGCACCAGCCAGATTCAGAGAGTGGCGCAGCAGCTGCTGGCTGGGGCTTCACCAACC 4277

Qy 1316 CCACACAGAGACACCAAGAGGCCCTTCTCTGTGGATGGGTACGAGCTGCACCCCG 1375
Db 4278 CCACACAGAGACACCAAGAGAGGCCCTTCTCTGTGGATGGGTACGAGCTGCACCCCG 4337
Qy 1376 ACAAGTGAACCGTGCAGCCCATCGAGCTGCCGAGAAAGAGAGCTGGACCGTGAACGACA 1435
Db 4338 ACAAGTGAACCGTGCAGCCCATCGTGTGCCGAGAAAGAGACAGCTGGACCGTGAACGACA 4397
Qy 1436 TCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCCGGCGATCAAGGTGC 1495
Db 4398 TCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCCGGCGATCAAGGTGC 4457
Qy 1496 GCCAGCTGTGCAAGCTGCTGGCGGCGCCAAAGGCCCTGACCGACATGCTGCTCCCTGACCG 1555
Db 4458 GCCAGCTGTGCAAGCTGCTGGCGGCGCCAAAGGCCCTGACCGAGTGTGCTCCCTGACCG 4517
Qy 1556 AGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTCGCGAGCCCGTGCACGGCG 1615
Db 4518 AGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTCAAGGAGCCCTGACCGCG 4577
Qy 1616 TGTACTAGGACCCAGCAAGACCTGGTGGCGGAGATCCAGAAAGCAGGGCCACGACCACT 1675
Db 4578 TGTACTAGGACCCAGCAAGACCTGGTGGCGGAGATCCAGAAAGCAGGGCCAGGGCCAGT 4637
Qy 1676 GGACCTACCGAGATCTACAGGAGCCCTTCAAGAACCTGAAAGACCGGCAAGTACGCAAGA 1735
Db 4638 GGACCTACCGAGATCTACAGGAGCCCTTCAAGAACCTGAAAGACCGGCAAGTACGCAAGA 4697
Qy 1736 TGCACACCGCCACACCAACACGCTGAAGCAGCTGACCGAGGCGGTGAGAAAGATCGGCA 1795
Db 4698 TGAAGGCGCCACACCAACACGCTGAAGCAGCTGACCGAGGCGGTGAGAAAGATCGGCA 4757
Qy 1796 TGGAGAGCATGCTGATCTGGGGGAGAGCCCAAGTTCGCGCTGCCATCCAGAAAGAGA 1855
Db 4758 CCGAGAGCATGCTGATCTGGGGGAGAGCCCAAGTTCGAGCTGCCATCCAGAAAGAGA 4817
Qy 1856 CCGAGAGCATGCTGAGGAGGAGTGGAGGAGTGGAGGAGTGGAGTGGAGTGGAGTGG 1915
Db 4818 CCGAGAGCATGCTGAGGAGGAGTGGAGGAGTGGAGGAGTGGAGTGGAGTGGAGTGG 4877
Qy 1916 TGAACACCCCGCTGCTGAGCTGAGTGTGTACCGAGTGGAGAGGAGGAGGAGGAGGAG 1975
Db 4878 TGAACACCCCGCTGCTGAGCTGAGTGTGTACCGAGTGGAGAGGAGGAGGAGGAGGAG 4937
Qy 1976 CCGAGAGCTTCTAGCTGGAGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCGCGCT 2035
Db 4938 CCGAGAGCTTCTAGCTGGAGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCGCGCT 4997
Qy 2036 ACGTGACCGAGCGGCGCGCGCAAGAGATCGTGAGCTGACCGAGACCAACCAACAGAGA 2095
Db 4998 ACGTGACCGAGCGGCGCGCGCAAGAGTGGTGGCTGACCGAGACCAACCAACAGAGA 5057
Qy 2096 CCGAGCTGAGGCGCATTCAGCTGGCCCTGCGAGGAGCGGCGAGGAGTGAACATCGTGA 2155
Db 5058 CCGAGCTGAGGCGCATTCAGCTGGCCCTGCGAGGAGCGGCGCTGAGAGTGAACATCGTGA 5117
Qy 2156 CCGAGAGCAGTACGCTTGGGCGATCTCCAGGCGCCAGCCGACAAAGAGCGAGAGCGAGC 2215
Db 5118 CCGAGAGCAGTACGCTTGGGCGATTCAGGCGCCAGCCGACAAAGAGCGAGAGCGAGC 5177
Qy 2216 TGGTGAACAGATCATCGAGAGCTGATCAAGAAAGAGAGTGTACCTGAGCTGGGTGC 2275
Db 5178 TGGTGAACAGATCATCGAGAGCTGATCAAGAAAGAGAGTGTACCTGAGCTGGGTGC 5237
Qy 2276 CCGGCCCAAGGGCATCGGCGCAACGAGCATCGAAGCTGGTGAAGAGGCGATCC 2335
Db 5238 CCGGCCCAAGGGCATCGGCGCAACGAGCATCGAAGCTGGTGAAGAGGCGATCC 5297
Qy 2336 GCAAGTGTCTTCTCTGGACGCGATCGATGGCGGATCGTGATGTACCAAGTA 2387
Db 5298 GCAAGTGTCTTCTCTGGACGCGATCGAAGGCCCGAGGAGGAGCGACGAGAA 5349

RESULT 11	AX427933	AX427933	9167 bp	DNA	linear	PAT 20-JUN-2002	718
LOCUS	Sequence 171 from Patent WO232943.						
DEFINITION	AX427933						
ACCESSION	AX427933						
VERSION	AX427933.1	GI:21538020					
KEYWORDS	synthetic construct.						
SOURCE	synthetic construct						
ORGANISM	artificial sequences.						
REFERENCE	1						
AUTHORS	Huang, Y. and Nabel, G. J.						
TITLE	Modifications of hiv env, gag, and pol enhance immunogenicity for genetic Immunization						
JOURNAL	Patent: WO 0232943-A 171 25-APR-2002;						
FEATURES	GOVERNMENT OF THE UNITED STATES (US)						
source	Location/Qualifiers						
	1..9167						
	/organism="synthetic construct"						
	/db_xref="taxon:32630"						
	/note="plasmid pYR1012x/s containing HIV genes"						
BASE COUNT	2225 a 2668 c 2517 g 1757 t						
ORIGIN							
Query Match	80.9%;	Score 1993.4;	DB 6;	Length 9167;			
Best Local Similarity	91.1%;	Pred. No. 2.1e-200;					
Matches 2182;	Conservative 0;	Mismatches 186;	Indels 27;	Gaps 5;			
Qy	14	TGGCGAGGCCATGAGCCAGG---CCACCAGCGCCCAACATCTGTGATGAGCGCAGCAACT	70				
Db	2964	TGGCGAGGCCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGAGCGCGCAACT	3023				
Qy	71	TCAGAGGCCCAAGCGCATCATCAAGTGTTCAACTGCGGCAAGAGAGGCCACATCGGCC	130				
Db	3024	TCCGCAACACGCGCAAGATCGTGAAGTGTCTCAACTGCGGCAAGAGAGGCCACACCGCC	3083				
Qy	131	GCAACTGCGCGCCCCCGCAAGAGGCGTCTGGAAGTGGCGCAAGAGAGGCCACACAGA	190				
Db	3084	GCAACTGCGCGCCCCCGCAAGAGGCGTCTGGAAGTGGCGCAAGAGAGGCCACACAGA	3143				
Qy	191	TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCGCGAGGAGCCTGGCCCTTCCGCCAGG	250				
Db	3144	TGAAGGACTGCACCGAGCGACAGGGCTAA-----TAGGGAAGATCTGSCCTTCCACAG	3197				
Qy	251	GCAAGGCCCGAGTTCCCGAGCGAGCAGACACCGCGCCCAACAGCCCCACAGCCCGGAGC	310				
Db	3198	GGAAGGCCAGGGAATTTTCTCAGAGCAGACAGAGCCCAACAGCCCCACCAAGAGAGAGC	3257				
Qy	311	TGCAGGTGCGGG-----CGACAACCCCCCGAGCGAGCGCGGCGCCAGCGCCAGGGCA	364				
Db	3258	TTCAAGTTTGGGAAGAGACACAACATCCCTCTCAGAGAGCAGGACCGATAGACAAAGAA	3317				
Qy	365	-----CCCTGAACTTCCCCAGATCACCTGTGGCAGCGCCCTTGGTGAGCATCAAGG	418				
Db	3318	CTGTATCTCTTTAGCTTCCCTCAGATCACTTTTGGCAGSCAGCCCTCGTCAACAATAAG	3377				
Qy	419	TGGCGGCGCAGATCAAGAGGCGCTGTGGACACCGCGCGCGAGCAGACCGTGTGGAGG	478				
Db	3378	TAGGGGCCAGCTGAAGAGGCCCTTCTAGACACCGGCGCGCGAGCACCCTGTGTGGAGG	3437				
Qy	479	AGATGAGCTGCCCGGAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCA	538				
Db	3438	AGATGAACCTGCCCGGCGCTGGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCA	3497				
Qy	539	AGGTGCGCCAGTAGACAGATCTCTGATGAGATCTGCGGCAAGAGGCCATCGGCACCG	598				
Db	3498	AGGTGGGCCAGTAGACAGATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG	3557				
Qy	599	TGCTGATCGGCCCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT	658				
Db	3558	TGCTGGTGGCCCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT	3617				

```
Db 4698 GCATGAAGGCGCCACACAGACGTGAAGCAGCTGACCGAGGCGGTGCACAAGATCG 4757
|||
QY 1793 CCATGAGAGCATCTGTGATCTGGGGCAAGACCCCAAGTTCGCCCTGCCCATCCAGAAG 1852
|||
Db 4758 CCACCGAGAGCATCTGTGATCTGGGGCAAGACCCCAAGTTCAAAGCTGCCCATCCAGAAG 4817
|||
QY 1853 AGACCTGGGAGACCTGGTGACCGACTACTTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1912
|||
Db 4818 AGACCTGGGAGGCGCTGGTGACCGAGTACTTGGCAGGCCACCTGGATCCCGAGTGGGAGT 4877
|||
QY 1913 TCCTGAACACCCCCCTGGTGGTGAAGCTGTGTACCAGCTGGAGAGAGGCCCATCATCG 1972
|||
Db 4878 TCCTGAACACCCCCCTGGTGGTGAAGCTGTGTACCAGCTGGAGAGAGGCCCATCATCG 4937
|||
QY 1973 GCGCCAGACCTTCTAGCTGGAGCGGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCG 2032
|||
Db 4938 GCGCCAGAGACCTTCTAGCTGGAGCGGCGCCCAACCGCGAGACCAAGCTGGGCAAGCGCG 4997
|||
QY 2033 GCTACGTGACCGACCGGGCGCGCAGAAAGATCGTGAGCCCTGACCGAGACCAACACAGCA 2092
|||
Db 4998 GCTACGTGACCGACCGGGCGCGCAGAAAGTGTGTGCGCCCTGACCGACACCAACACAGCA 5057
|||
QY 2093 AGACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACAGCGCGAGGTGAACATCG 2152
|||
Db 5058 AGACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACAGCGCGCTGGAGGTGAACATCG 5117
|||
QY 2153 TGACCGACAGCCAGTAGCCCTGGGCATCATCCAGGCCCGCCAGCAGAGCGAGAGCG 2212
|||
Db 5118 TGACCGACAGCCAGTAGCCCTGGGCATCATCCAGGCCCGCCAGCAGAGCGAGAGCG 5177
|||
QY 2213 AGCTGTGTAACCAAGATCATCGACGAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 2272
|||
Db 5178 AGCTGTGACCCAGATCATCGACGAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 5237
|||
QY 2273 TGCCCGCCCAAGAGGATCGCGCGCAACAGCAGATCGACAGAGCTGGTGAGCAAGGGCA 2332
|||
Db 5238 TGCCCGCCCAAGAGGATCGCGCGCAACAGCAGAGGTGGACGCGCTGGTGAGCGCGGCA 5297
|||
QY 2333 TCCGCAAGGTGCTTCTTCCGCGGATCATCGCGCGCTGATCTACCAAGTGA 2387
|||
Db 5298 TCCGCAAGGTGCTTCTTCCGCGGATCATCGACAAAGGCCCGCAGGAGGACGACAGAA 5352
|||

RESULT 12
AX427928
LOCUS AX427928 9170 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 166 from Patent W00232943.
ACCESSION AX427928
VERSION AX427928.1 GI:21538015
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial construct.
REFERENCE
1 Huang, Y. and Nabel, G.J.
Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 166 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
source 1..9170
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="plasmid pVR1012x/s containing HIV genes"

BASE COUNT 2225 a 2669 c 2519 g 1757 t
ORIGIN
Query Match 80.9%; Score 1993.4; DB 6; Length 9170;
Best Local Similarity 91.1%; Pred. No. 2.1e-200;
Matches 2182; Conservative 0; Mismatches 186; Indels 27; Gaps 5;
QY 14 TGGCCGAGGCCATGAGCCAGG---CCACCAGCGCCAAACATCTCTGATGTCAGCGGCAACT 70
|||||
```

```
Db 2967 TGGCCGAGGCGCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGTCAGCGGCAACT 3026
|||||
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTCTTCAACTGCGGCAAGGAGGGCCACATCGCCC 130
|||
Db 3027 TCCGCAACCAAGCGCAAGATCGTGAAGTGTCTCAACTGCGGCAAGGAGGGCCACACGCC 3086
|||||
QY 131 GCAACTGCGCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA 190
|||||
Db 3087 GCAACTGCGCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA 3146
|||||
QY 191 TGAAGGACTGCAACCGAGCGCCCAACCTTCTTCCGCGAGGACCTGGCCCTTCCCCCAGG 250
|||||
Db 3147 TGAAGGACTGCAACCGAGCGACAGGCTAA-----TAGGGAAGATCTGCGCTTCCCCAAG 3200
|||||
QY 251 GCAAGGCCCGCGAGTTCCTCCCGAGCAGCAGAGAACCGCGCCCAACAGCCCCACCGCGCGAGC 310
|||||
Db 3201 GGAAGGCCAGGGAATTTTCTTCAGAGCAGACAGAGCCAAAGAGCCCCCACCAGAAGAGAGC 3260
|||||
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGAGCGCGCAGGCGCA 364
|||||
Db 3261 TTCAAGTTTGGGGAAGAGACAACTCCCTCTCAGAAAGCAGGAGCGGATAGACAAGGAA 3320
|||||
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTGGTGAGCATCAAGG 418
|||||
Db 3321 CTGTATCTCTTAGCTTCCCTCAGATCACTCTTGGCAGCGACCCCTCGTCACATAAAGA 3380
|||||
QY 419 TGGCGCGCAGATCAAGGAGGCGCTCTGTGACACCGCGCGCAGCAGACCGCTGTGGAGG 478
|||||
Db 3381 TAGGGGCCAGCTGAAGAGGAGGCGCTTCTAGACACCGCGCGCAGCAGCAGCTGTGGAGG 3440
|||||
QY 479 AGATGAGCTTSCCGCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
|||||
Db 3441 AGATGAACCTGCGCGCGCGCTGGAAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 3500
|||||
QY 539 AGTGCGCCAGTAGGAGCAGATCTGTATCAGATCTCTCGGCAAGAGAGGCGCATCGGACCG 598
|||||
Db 3501 AGGTGGGCCAGTAGGAGCAGATCTGTATCAGATCTCTCGGCGCACAAAGGCCATCGGACCG 3560
|||||
QY 599 TGTGTATCGCGCGCCACCGCGTGAACATCATCGCGCGCAACATGCTCACCGCTGGCT 658
|||||
Db 3561 TGTGTGGCGCGCCACCGCGTGAACATCATCGCGCGCAACATGCTCACCGCTGGCT 3620
|||||
QY 659 GCACCTGAACTTCCCGCATCAGCCCGCATCGAGACCGCTGCGCGTGAAGCTGAAGCGCGCA 718
|||||
Db 3621 GCACCTGAACTTCCCGCATCAGCCCGCATCGAGACCGCTGCGCGTGAAGCTGAAGCGCGCA 3680
|||||
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGTACCGGAGGAGAGATCAAGGCCCTGACCG 778
|||||
Db 3681 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGTACCGGAGGAGAGATCAAGGCCCTGCTGG 3740
|||||
QY 779 CCATCTGCGAGGAGATGGAGAAGGCGCAAGATCACCAAGATCGCGCGCGAGAACCCCT 838
|||||
Db 3741 AGATCTGCACCGAGATGGAGAAGGCGCAAGATCACCAAGATCGCGCGCGAGAACCCCT 3800
|||||
QY 839 ACAACACCCCGCTTTCGCCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
|||||
Db 3801 ACAACACCCCGCTTTCGCCCATCAAGAAGAGGAGACAGCACCAAGTGGCGCAAGCTGGTGG 3860
|||||
QY 899 ACTTCCGCGAGCTGAACAAGCGCACCGAGACTTCTGGGAGGTGCAGCTGGGCATCCCC 958
|||||
Db 3861 ACTTCCGCGAGCTGAACAAGCGCACCGAGACTTCTGGGAGGTGCAGCTGGGCATCCCC 3920
|||||
QY 959 ACCCGCGCGCTTGAAGAAGAGAGCGTGAACCGTGTGGAGCTGGGCGACCGCTACT 1018
|||||
Db 3921 ACCCGCGCGCTTGAAGAAGAGAGCGTGAACCGTGTGGAGCTGGGCGACCGCTACT 3980
|||||
QY 1019 TCAGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 1078
|||||
Db 3981 TCAGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 4040
|||||
QY 1079 ACAACGAGACCCCGGATCCGCTACCAAGTACACAGTGTCTGCGCGCAGGCTGGAGGCA 1138
|||||
```

Db 4041 ACAACGAGACCCCGGCATCCGCTACCAAGTACAACGTCTGCCCGCAGGGCTGGAAGGCA 4100
QY 1139 GCCCGACATCTTCCAGAGCAGCATGACCAAGATCCTTGGAGCCCTTCGCGCCCGCAACC 1198
Db 4101 GCCCGCATCTTCCAGTGACAGATGACCAAGATCCTTGGAGCCCTTCCGCAAGCAGAAC 4160
QY 1199 CCGAGATCGTATACCA-----GCCCCCTTGTAGTGGGAGGACCTGGAGATCG 1252
Db 4161 CCGACATCGTATACCAAGTACATGACCACTGTACGTGGGAGGACCTGGAGATCG 4220
QY 1253 GCCAGCAGCGCCCAAGATCAGAGATGCGCAAGCAGCCTCTCGGCTGGGGCTTCACCA 1312
Db 4221 GCCAGCAGCGCCCAAGATCAGAGATGCGCAAGCAGCCTCTCGGCTGGGGCTTCACCA 4280
QY 1313 CCCCAGCAAGACACAGAGAGGAGCCCTTCTCTGTGTGATGGCTTACAGCTGCACC 1372
Db 4281 CCCCAGCAAGACACAGAGAGGAGCCCTTCTCTGTGTGATGGCTTACAGCTGCACC 4340
QY 1373 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGTGGACCTGGAACG 1432
Db 4341 CCGACAAGTGGACCGTGCAGCCCATCGTGTGCCGAGAGAGAGTGGACCTGGAACG 4400
QY 1433 ACATCCAGAGCTGTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGCATCAAG 1492
Db 4401 ACATCCAGAGCTGTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGCATCAAG 4460
QY 1493 TGCGCCAGCTGTGAAGTGTGCGCGCGCCCAAGGCGCTGACCGACATCTGTCGCCCTGA 1552
Db 4461 TGCGCCAGCTGTGAAGTGTGCGCGCGCCCAAGGCGCTGACCGAGTGTGTCGCCCTGA 4520
QY 1553 CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGGGAGCCCTGTGCACG 1612
Db 4521 CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGAAGGAGCCCTGTGCACG 4580
QY 1613 GCCTGTACTAGCACCCAGCAAGCACTGTGTGGCGGAGATCCAGAGCAGGCGCCACGAC 1672
Db 4581 GCCTGTACTAGCACCCAGCAAGCACTGTGTGGCGGAGATCCAGAGCAGGCGCCAGGCGC 4640
QY 1673 AGTGGACCTACCAAGTCTACCAAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCA 1732
Db 4641 AGTGGACCTACCAAGTCTACCAAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCA 4700
QY 1733 AGATGGCACCAGCCACCAACAGCTGAAGCAGCTGACCGAGGCGGTGCAGAGATCG 1792
Db 4701 GCATGAAGGCGGCCACCAACAGCTGAAGCAGCTGACCGAGGCGGTGCAGAGATCG 4760
QY 1793 CCATGGAGAGCATCGTGTATCTGGGCGAAGCCGCCAAGTTCGGCTCGCCCATCCAGAGG 1852
Db 4761 CCACCGAGAGCATCGTGTATCTGGGCGAAGCCGCCAAGTTCAGCTGCCCATCCAGAGG 4820
QY 1853 AGACCTGGGAGACCTGTGGACCGACTACTTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1912
Db 4821 AGACCTGGGAGGCTGTGTGACCGAGTACTTGGCAGGCCACCTGGATCCCGAGTGGGAGT 4880
QY 1913 TCGTGAACACCCCCCTGTGTGAGCTGTGTACAGCTGAGAGAGGAGCCCATCATCG 1972
Db 4881 TCGTGAACACCCCCCTGTGTGAGCTGTGTACAGCTGAGAGAGGAGCCCATCATCG 4940
QY 1973 GCGCGGAGAGCTTCTAGCTGGACCGCGCCCAACCGCAGACCAAGATCGGCAAGGCGC 2032
Db 4941 GCGCGGAGAGCTTCTAGCTGGACCGCGCCCAACCGCAGACCAAGATCGGCAAGGCGC 5000
QY 2033 GCTACGTGACCGAGCGCGGCGGAGAGATCGTGTAGCCTGACCGAGACCAACCAAGCA 2092
Db 5001 GCTACGTGACCGAGCGCGGCGGAGAGTGTGTGCGCTGACCGAGACCAACCAAGCA 5060
QY 2093 AGACCGAGCTGCAGGCCATCCAGCTGGCTGTGAGGACAGGCGGAGAGGTGAACATCG 2152
Db 5061 AGACCGAGCTGCAGGCCATCCAGCTGGCTGTGAGGACAGGCGGCTGTGAGGTGAACATCG 5120
QY 2153 TGACCCGACGAGTACGCGCTGGGCATCATCCAGGCGCCAGCGGCGGAGAGGCG 2212
Db 5121 TGACCCGACGAGTACGCGCTGGGCATCATCCAGGCGCCAGCGGCGGAGAGGCG 5180

QY 2213 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGG 2272
Db 5181 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGG 5240
QY 2273 TGCCCGCCCAAGAAGGAGTTCGGCGGCAACGAGCAGATCGAAGCTGGTGAAGAGGCA 2332
Db 5241 TGCCCGCCCAAGAAGGAGTTCGGCGGCAACGAGCAGTGGAGCGCTGGTGAAGCGGCA 5300
QY 2333 TCCGCAAGTGTCTTCTCTGGACGCGATCGATGGCGCATCTGTATCTACCACTA 2387
Db 5301 TCCGCAAGTGTCTTCTCTGGACGCGATCGAAGCGCCAGGAGGAGCAGCAAA 5355

RESULT 13
AX427935
LOCUS AX427935 9782 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 173 from Patent WO0232943.
ACCESSION AX427935
VERSION AX427935.1 GI:21538022
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 173 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
source 1..9782
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2339 a 2895 c 2719 g 1829 t
ORIGIN

Query Match 80.9%; Score 1993.4; DB 6; Length 9782;
Best Local Similarity 91.18; Pred. No. 2.1e-200;
Matches 2182; Conservative 0; Mismatches 186; Indels 27; Gaps 5;

QY 14 TGGCGGAGGCGCATGAGCCAGG---CCACCAGCGCAACATCTGTATGAGGCGCAGCAACT 70
Db 3579 TGGCGGAGGCGCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGAGCGGCAACT 3638
QY 71 TCAGAGGCGCCCAAGCGCATCATCAAGTCTTCAACTGGCGCAAGAGGCGCACATCGGCC 130
Db 3639 TCCGCAACCCAGCGCAAGATCGTGAAGTCTTCAACTGGCGCAAGAGGCGCACACCGCC 3698
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGAGGCGCCACCAAGA 190
Db 3699 GCAACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGAGGCGCCACCAAGA 3758
QY 191 TGAAGGACTGCACCGAGCGCCAGGCGCAACTTCTTCCGCGAGGAGCTGCGCTTCCCGCAGG 250
Db 3759 TGAAGGACTGCACCGAGCGCGACAGGCGTAA-----TAGGGAAGATCTGCGCTTCCCAACAG 3812
QY 251 GCAAGGCGCGCGAGTTCCTCCAGCGCAGAGACCGCGCCCAACAGCCCCACAGCGCGGAGC 310
Db 3813 GGAAGGCGCGGAGATTTTCTTTCAGAGCAGACAGACAGCGCCCAACAGCCCCACAGAGAGAGC 3872
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCGCGCA 364
Db 3873 TTCAGTTTGGGGAAGAGAGACAACTCCCTCTCAGAGAGAGGAGCCCATAGACAAAGAA 3932
QY 365 -----CCCTGAACCTTCCCCAGATCACTCTGTGGCAGCGCCCTTGTGTGAGCATCAAG 418
Db 3933 CTGTATCTTTAGCTTCCCTCAGATCACTTGTGGCAGCGACCCCTCTCTCACAATAAGA 3992
QY 419 TGGCGCGCCAGATCAAGAGGCGCCCTGTGTGACACCGCGCGCGCGCGCGCGCGCGTGTGAGG 478
Db 3993 TAGGGGCGCAGCTGAAGGAGGCGCCCTTCTAGACACCGCGCGCGCGCGCGCGCGCGTGTGAGG 4052

FEATURES		Location/Qualifiers	
source		1..9783	
		/organism="synthetic construct"	
		/db_xref="taxon:32630"	
		/note="plasmid pVR1012x/s containing HIV genes"	
BASE COUNT	2339 a	2897 c	2723 g 1824 t
ORIGIN			
Query Match	80.9%;	Score 1993.4;	DB 6; Length 9783;
Best Local Similarity	91.1%;	Pred. No. 2.1e-200;	
Matches 2182;	Conservative 0;	Mismatches 186;	Indels 27; Gaps 5;
QY	14	TGGCGGAGGCATGAGCCAGG-----CCACCAAGGCGCAACATCTGTATGACGCGCAGCAACT	70
Db	2964	TGGCGGAGGCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGACGCGCGCAACT	3023
QY	71	TCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATCGGCC	130
Db	3024	TCCGCAACCAAGCGCAAGATCTGTGAAGTGTCTCAACTGCGGCAAGGAGGGCCACACGGCC	3083
QY	131	GCAACTGCCCGCCCCCGCCGAAGAGGGCTGTGTGAAGTGGCGCAAGGAGGGCCACACAGA	190
Db	3084	GCAACTGCCCGCCCCCGCCGAAGAGGGCTGCTGTGAAGTGGCGCAAGGAGGGCCACACAGA	3143
QY	191	TGAAGGACTTCACCGAGCGCCAGGCCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCGAGG	250
Db	3144	TGAAGGACTTCACCGAGCGCCAGGCGTAA-----TAGGGAAGATCTGGCCTTCCCAACAAG	3197
QY	251	GCAAGGCCCGGAGTTCCCGAGGAGAGAACCGCGCCCAACAGCCCCACAGCCCGCGAGC	310
Db	3198	GGAAGGCCGAGGAATTTTCTTCAGAGGAGACCAAGAGCCCAACAGCCCCACCAAGAGAGAGC	3257
QY	311	TGCAGGTGCGGGG-----CGACAAACCCCGCAGCGAGGCGCGCGCGAGCGCCAGGGCA	364
Db	3258	TTCAAGTTTGGGAGAGACACAACACTCCCTCTCAGAAAGCAGGAGCGCGATAGACAAGAA	3317
QY	365	-----CCCTGAACTTCCCGCAGATCACCTGTGTGCAGCGGCCCTTGGTGAAGCATCAAGG	418
Db	3318	CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAACAATAAGA	3377
QY	419	TGGCGGCGCAGATCAAGGAGGCCCTGTGTGACACCGCGCGCGAGCACACCGTGTGGAGG	478
Db	3378	TAGGGGGCAGCTGAAGAGGGCCCTTTAGACACCGCGCGCGAGCACACCGTGTGGAGG	3437
QY	479	AGATGAGCTTCCCGGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGGCTTTCATCA	538
Db	3438	AGATGAACTTGCCCGGCGCTTGGAAAGCCCAAGATGATCGCGGCGATCGCGGCTTTCATCA	3497
QY	539	AGGTGCGCCAGTAGACACAGATCTCTGTATCGAGATCTGGGCAAGAGGCCATCGGCACCG	598
Db	3498	AGGTGGGCGAGTAGACACAGATCTCTGTATCGAGATCTGGGCGCACAAAGGCGCATCGGCACCG	3557
QY	599	TGCTGATCGGCCCCACCCCGTGAACATCATCGGCCCGCAACATGCTGACCCAGTGGGCT	658
Db	3558	TGCTGGTGGGCGCCACCCCGTGAACATCATCGGCCCGCAACCTGTGACCCAGATCGGCT	3617
QY	659	GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGGCCGCTGAAGTGAAGCCCGGCA	718
Db	3618	GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGGCCGCTGAAGTGAAGCCCGGCA	3677
QY	719	TGGAGCGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG	778
Db	3678	TGGAGCGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTTGGTGG	3737
QY	779	CCATCTGGGAGGAGATGGAGAGGAGGCGCAAGATCACCAAGATCGGCCCGGAGAACCCCT	838
Db	3738	AGATCTGCCGAGAGATGGAGAGGAGGCGCAAGATCAGCAAGATCGGCCCGGAGAACCCCT	3797
QY	839	ACAACACCCCGTGTTCGCCATCAAGAAAGAGGACACCAAGTGGCGCAAGCTGGTGG	898
Db	3798	ACAACACCCCGTGTTCGCCATCAAGAAAGAGGACACCAAGTGGCGCAAGCTGGTGG	3857
QY	899	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGGAGGTGAGCTGGGATCCCCC	958
Db	3858	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTGAGCTGGGATCCCCC	3917
QY	959	ACCCCGCGCGCTCAAGAAAGAGAGCGGTGACCGTGTGAGAGTGGCGGACCGCTACT	1018
Db	3918	ACCCCGCGCGCTCAAGCAAGAGAGCGGTGACCGTGTGAGAGTGGCGGACCGCTACT	3977
QY	1019	TCAGCGTCCCGCTGGAGAGGACTTCCGCAAGTACACCGCTTTCACATCCCGAGCATCA	1078
Db	3978	TCAGCGTCCCGCTGGACAAGGACTTCCGCAAGTACACCGCTTTCACATCCCGAGCATCA	4037
QY	1079	ACAACGAGACCCCGGCATCCGCTACCACTACCAAGTGTGCTGCCAGGCGTGGAAAGGCA	1138
Db	4038	ACAACGAGACCCCGGCATCCGCTACCACTACCAAGTGTGCTGCCAGGCGTGGAAAGGCA	4097
QY	1139	GCCCCACATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTTCGCGCGCCGCAACC	1198
Db	4098	GCCCCACATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCCCTTTCGCGCAACGACACC	4157
QY	1199	CCGAGATCTGTATCTACCA-----GGCCCCCTGTACGTGGGCGAGGACCTGGAGATCG	1252
Db	4158	CCGACATCTGTATCTACCAAGTACATGGACACCTGTACGTGGGCGAGGACCTGGAGATCG	4217
QY	1253	GCCAGCACCGCGCCCAAGATCGAGGAGCTGGCAAGCACCTGCTGCGTGGGCGTTCACCA	1312
Db	4218	GCCAGCACCGCGCCCAAGATCGAGGAGCTGGCCAGCACCTGCTGCGTGGGCGTTCACCA	4277
QY	1313	CCCCCGACAAAGACACCAAGAGGAGCCCGCTTCTGTGTGATGGGTACGAGCTGCACC	1372
Db	4278	CCCCCGACAAAGACACCAAGAGGAGCCCGCTTCTGTGTGATGGGTACGAGCTGCACC	4337
QY	1373	CCGCAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG	1432
Db	4338	CCGCAAGTGGACCGTGCAGCCCATCGTGTGCCCGCAGAAAGACAGCTGGAACGTTGAACG	4397
QY	1433	ACATCCAGAAAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCCAGATCTACCCCGGCAATCAAG	1492
Db	4398	ACATCCAGAAAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCCAGATCTACCCCGGCAATCAAG	4457
QY	1493	TGCGCCAGCTGTGCAAGCTGTGCGCGCGGCCAAAGGCCCTGACCGACATCTGTCGCCCTGA	1552
Db	4458	TGCGCCAGCTGTGCAAGCTGTGCGCGCGGCCAAAGGCCCTGACCGAGGTGTGTCGCCCTGA	4517
QY	1553	CCGAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGCGGAGCCCGTGCACG	1612
Db	4518	CCGAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGAGGAGCCCGTGCACG	4577
QY	1613	GGGTGTACTAGACCCCGACAGGAGCTGGTGGCCGAGATCCAGAGCAGGCGCCACGAGCC	1672
Db	4578	GGGTGTACTAGACCCCGACAGGAGCTGTATCGCCGAGATCCAGAGCAGGCGCCAGGGCC	4637
QY	1673	AGTGGACCTACCATCTTACCGAGGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA	1732
Db	4638	AGTGGACCTACCATCTTACCGAGGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA	4697
QY	1733	AGATGCGCACCGCCCGACACCAACGACGCTGAAGCAGCTGACCGAGGCGCGTGCAGAAAGTCG	1792
Db	4698	GCATGAAGGCGCGCCACACCAACGACGCTGAAGCAGCTGACCGAGGCGCGTGCAGAAAGTCG	4757
QY	1793	CCATGGAGAGCATCTGTATCTTGGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAAG	1852
Db	4758	CCACCGAGAGCATCTGTATCTTGGGGCAAGACCCCAAGTTCAGAGTGCCTCATCCAGAAAG	4817
QY	1853	AGACCTGGGAGACCTGGTGGACCGACTTCTGGCAGGCCACCTGGATCCCGAGTGGGAGT	1912
Db	4818	AGACCTGGGAGGCGCTGGTGGACCGAGTACTTGGCAGGCCACCTGGATCCCGAGTGGGAGT	4877
QY	1913	TGCTGAACACCCCGCTTGGTGAAGCTGTGTACAGCTGGAGAGGAGGCGCCATCATCG	1972
Db	4878	TGCTGAACACCCCGCTTGGTGAAGCTGTGTACAGCTGGAGAGGAGGCGCCATCATCG	4937
QY	1973	GCGCCGAGACCTTCTAGCTGGAGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCCG	2032

Db 4938 GCGCGAGACCTTCTACGTGGACGGCGCGCCAAACCGCGAGACCAAGCTGGGCAAGCGCG 4997
Qy 2033 GCTACGTACCGACCGGGCGGCGGAGAGATCGTGAGCCTGACCGAGACCAACACAGCA 2092
Db 4998 GCTACGTACCGACCGGGCGGCGGAGAGTGTGTCCTTGACCGAGACCAACACAGCA 5057
Qy 2093 AGACCGAGTGCAGGCGCATCCAGTGTGCGCCTGTCAGGACACGGCGAGGTGAACATCG 2152
Db 5058 AGACCGAGTGCAGGCGCATCCAGTGTGCGCCTGTCAGGACACGGCGGTGAGGTGAACATCG 5117
Qy 2153 TGACCGAGACGAGTACGCGCTGGGCATCATCCAGGCGCCAGCCGACAGAGCGAGCG 2212
Db 5118 TGACCGAGACGAGTACGCGCTGGGCATCATCCAGGCGCCAGCCGACAGAGCGAGCG 5177
Qy 2213 AGTGTGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGATCTGAGCTGGG 2272
Db 5178 AGTGTGTGAGCCAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGATCTGAGCTGGG 5237
Qy 2273 TGCCCGCCACAGAGGCGATCGGGGCGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGCGCA 2332
Db 5238 TGCCCGCCACAGAGGCGATCGGGGCGCAACGAGCAGGTGGACGGCCTGGTGAGCGCGCGCA 5297
Qy 2333 TCCGCAAGGTGCTTCTTCTGAGGCGCATCGATGCGGCGCATCGTGTATACAGTA 2387
Db 5298 TCCGCAAGGTGCTTCTTCTGAGCGCATCGACAAAGGCCCGAGGAGCAGAGAA 5352

RESULT 15
AX427932
LOCUS AX427932 9792 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 170 from Patent W00232943.
ACCESSION AX427932
VERSION AX427932.1 GI:21538019
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 170 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source 1.
1. .9792
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"

BASE COUNT 2339 a 2898 c 2725 g 1830 t
ORIGIN

Query Match 80.9%; Score 1993.4; DB 6; Length 9792;
Best Local Similarity 91.1%; Pred. No. 2.1e-200;
Matches 2182; Conservative 0; Mismatches 186; Indels 27; Gaps 5;

Qy 14 TGCGCGAGGCGCATGAGCGAGG---CCACAGCGCCACATCTCTGATGACGCGCAGCAACT 70
Db 2967 TGCGCGAGGCGCATGAGCGAGGTCACCAACAGCGCCACCATCATGATGACGCGGCAACT 3026
Qy 71 TCAAGGCGCCCAAGCGCATCATCAAGTGTTCACACTGCGCGAGGAGGCGGCACATCGCCC 130
Db 3027 TCGCGAACACGCGCAGATCGTGAAGTGTTCACACTGCGCGAGGAGGCGGCACACCGCCC 3086
Qy 131 GCAACTGCGCGCGCCCGCGCAAGAAGGCTGCTGGAAGTGGCGCAAGAGGCGGCACACAGA 190
Db 3087 GCAACTGCGCGCGCCCGCGCAAGAAGGCTGCTGGAAGTGGCGCAAGAGGCGGCACACAGA 3146
Qy 191 TGAAGACTGACCGAGCGCCAGGCGCAACTTCTTCGCGAGGAGACCTTGGCCCTTCCCGCAGG 250
Db 3147 TGAAGACTGACCGAGCGCGCAGCGAGTAA-----TAGGGAAGATCTGGCCCTTCCCAACAAG 3200
Qy 251 GCAAGGCGCGGAGTTCCTCCAGCGGAGCAGACCGCGCGCAACCGCGCGACCCACCGCGCGAGC 310
Db 4281 CCCCCGACAGAAGACACCAAGAGGAGCCCCCTTCTCTGTGGATGGGCTACGAGCTGCACC 4340

Db 3201 GGAAGCGAGCGGAATTTTCTTACAGAGCAGACCAAGAGCCACAGCCCGACCAAGAGAGC 3260
Qy 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGAGCGCCAGAGGCA 364
Db 3261 TTCAGGTTTGGGGAAGAGACAACAACCTCCCTCTCAGAAGCAGGAGCGGATAGACAAGGAA 3320
Qy 365 -----CCCTGAACTTCCCCCAGATCACCTCTGTGGCAGCGCCCTGTGTGAGCATCAAG 418
Db 3321 CTGTATCTTCTTAGCTTCCCTCAGATCACTCTTTGCGAGCGACCCCTCGTCAATAAAGA 3380
Qy 419 TGGCGGCGCAGATCAAGAGAGCCCTGTGTGACACCGCGCGCGAGACACCTGTGTGAGG 478
Db 3381 TAGGGGCGCAGCTGAAGAGAGGCCCTTCTAGACACCGCGCGCGAGACACCTGTGTGAGG 3440
Qy 479 AGATGAGCCTGCGCGGCAAGTGGAAAGCCCAAGATGATCGGCGGATCGGCGGCTTCATCA 538
Db 3441 AGATGAGACCTGCGCGCGCGCTGGAAAGCCCAAGATGATCGGCGGATCGGCGGCTTCATCA 3500
Qy 539 AGGTGCGCCAGTAGTACGACAGATCCTGTATCGAGATCTTCGGCAAGAAGGCCATCGGCACCG 598
Db 3501 AGGTGCGCCAGTAGTACGACAGATCCTGTATCGAGATCTTCGGCCACAAGGCCATCGGCACCG 3560
Qy 599 TGCTGATCGGCGCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT 658
Db 3561 TGCTGATCGGCGCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT 3620
Qy 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGCCTGAAAGTGAAGTGAAGCGCGCA 718
Db 3621 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGCCTGAAAGTGAAGTGAAGCGCGCA 3680
Qy 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGGAGAGAGATCAAGGCCCTTGACCG 778
Db 3681 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGGAGAGAGATCAAGGCCCTTGAGTG 3740
Qy 779 CCATCTCGGAGGAGATGGAAAGAGGCGCAAGATCACCAGAGATCGGCGCGCGAGAACCCCT 838
Db 3741 AGATCTGCACCGAGATGGAAAGAGGCGCAAGATCACCAGAGATCGGCGCGCGAGAACCCCT 3800
Qy 839 ACAACACCCCGCTGTTCGCCATCAAGAAGAGGACACACCAAGTGGCGCAAGCTGGTGG 898
Db 3801 ACAACACCCCGCTGTTCGCCATCAAGAAGAGGACACACCAAGTGGCGCAAGCTGGTGG 3860
Qy 899 ACTTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGTGGGCAATCCCC 958
Db 3861 ACTTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGTGGGCAATCCCC 3920
Qy 959 ACCCGCGCGCTGAAGAAGAGAGCGTGCCTGAGAGTGGCGCGCGAGACCCCTACT 1018
Db 3921 ACCCGCGCGCTGAAGAAGAGAGCGTGCCTGAGAGTGGCGCGCGAGACCCCTACT 3980
Qy 1019 TCAGCGTGCCTTGGAGAGGACTTCCGCAAGTACACCGCTTCAACCATCCCCAGCATCA 1078
Db 3981 TCAGCGTGCCTTGGAGAGGACTTCCGCAAGTACACCGCTTCAACCATCCCCAGCATCA 4040
Qy 1079 ACAACGAGACCCCGGATCCGCTACCAAGTACAAAGTGTGCTGCCCCAGGGGTGGAAGGCA 1138
Db 4041 ACAACGAGACCCCGGATCCGCTACCAAGTACAAAGTGTGCTGCCCCAGGGGTGGAAGGCA 4100
Qy 1139 GCGCGAGCATCTTCCAGAGCAGCATGACCAAGATCTTGGAGCCCTTCCCGCGCGCGCAACC 1198
Db 4101 GCGCGAGCATCTTCCAGTGCAGCATGACCAAGATCTTGGAGCCCTTCCCGCGCGCGCAACC 4160
Qy 1199 CCAGAGTCTGTATCTACCA-----GGCCCCCTGTACGTGGGCGAGGAGCTGGAGATCG 1252
Db 4161 CCAGATCTGTATCTACCAAGTACATGACCAACCTGTACGTGGGCGAGGAGCTGGAGATCG 4220
Qy 1253 GCAGACACCGCGCAGATCGAGAGCTGCGCAAGCAACCTGCTGCGGTGGGGGTTCACCA 1312
Db 4221 GCAGACACCGCGCAGATCGAGAGCTGCGCAAGCAACCTGCTGCGGTGGGGGTTCACCA 4280
Qy 1313 CCCCCGACAGAAGACCAAGAGGAGCCCCCTTCTCTGTGGATGGGCTAGGAGTGCACC 1372
Db 4281 CCCCCGACAGAAGACCAAGAGGAGCCCCCTTCTCTGTGGATGGGCTAGGAGTGCACC 4340

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 10:51:03 ; Search time 491 seconds
(without alignments)
11296.684 Million cell updates/sec

Title: US-09-610-313-31

Perfect score: 2463

Sequence: 1 gtcagccaccatggccga.....gggtagcaccggtgaattc 2463

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

	1:	2:	3:	4:	5:	6:	7:	8:	9:	10:	11:	12:	13:	14:	15:	16:	17:	18:	19:	20:	21:	22:	23:	24:
	/SID22/gcgdata/geneseq/geneq-emb1/NA1980.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1981.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1982.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1983.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1984.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1985.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1986.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1987.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1988.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1989.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1990.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1991.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1992.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1993.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1994.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1995.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1996.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1997.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1998.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA1999.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA2000.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA2001A.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA2001B.DAT.*	/SID22/gcgdata/geneq/geneq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2463	100.0	2463	24	ABL39960
2	2442.2	99.2	2469	24	ABL39959
3	2436.2	98.9	2457	24	ABL39961
4	2046	83.1	2306	21	AA70479
5	2027.6	82.3	9166	24	ABK91616
6	2025.2	82.2	2312	21	AAA70481
7	2019.2	82.0	2300	21	AAA70480
8	2019.2	82.0	9788	24	ABK91622
9	2005.2	81.4	9169	24	ABK91617

10	2001.8	81.3	9194	24	ABK91612	Modified HIV prote
11	2000.2	81.2	9194	24	ABK91611	Modified HIV prote
12	2000.2	81.2	12411	24	ABK91613	Modified HIV prote
13	1996.8	81.1	9785	24	ABK91624	Modified HIV prote
14	1993.4	80.9	9167	24	ABK91619	Modified HIV prote
15	1993.4	80.9	9170	24	ABK91614	Modified HIV prote
16	1993.4	80.9	9782	24	ABK91621	Modified HIV prote
17	1993.4	80.9	9783	24	ABK91620	Modified HIV prote
18	1993.4	80.9	9792	24	ABK91618	Modified HIV prote
19	1991.8	80.9	9189	24	ABK91607	Modified HIV prote
20	1972.2	80.1	3015	24	ABL39983	Synthetic Pol poly
21	1955.6	79.4	3009	24	ABL40024	HIV Gag-polymerase
22	1942.6	78.9	4319	21	AAA70414	Packaging construc
23	1909	77.5	8908	21	AAZ52055	HIV-1 subtype C is
24	1886.6	76.6	2577	24	AAAL1592	Nucleotide sequenc
25	1879	76.3	4343	24	AAI67896	HIV FS(+)-ProtInac
26	1878.8	76.3	2305	21	AAA70477	Nucleotide sequenc
27	1875	75.1	4341	24	AAI67897	Synthetic protease
28	1857.8	75.4	1965	24	ABL39989	Synthetic protease
29	1854.4	75.3	1978	24	ABL39991	Modified HIV prote
30	1852.4	75.2	7897	24	ABK91609	HIV FS(+)-ProtInac
31	1852	75.2	2299	21	AAA70478	Codon optimised Hu
32	1788.8	72.6	3012	21	AAZ52054	Modified HIV prote
33	1784	72.4	9407	20	ABK91623	Modified HIV prote
34	1632.6	66.3	4307	20	AAZ08740	HIV gagpol-SYNGP c
35	1632.6	66.3	4307	21	AAA93972	Human immunodefici
36	1632.6	66.3	4307	22	AAC86876	Nucleotide sequenc
37	1624.6	66.0	4307	22	AAH43696	Codon optimised HI
38	1624.6	66.0	4327	21	AAA93984	HIV partial leader
39	1624.6	66.0	4353	21	AAA93983	HIV partial leader
40	1624.6	66.0	4642	21	AAA93982	HIV complete leade
41	1624.6	66.0	9772	22	AAH43699	Codon optimised ga
42	1581.2	64.2	1680	24	ABL39997	Synthetic RT polyn
43	1575.2	64.0	1668	24	ABL39999	Synthetic RT polyn
44	1566	63.6	4338	22	AAH22806	DNA sequence of a
45	1566	63.6	8366	22	AAH22810	DNA sequence of a

ALIGNMENTS

RESULT 1
ABL39960
ID ABL39960 standard; DNA; 2463 BP.
XX
XX ABL39960;
XX
XX
DT 15-MAY-2002 (first entry)
XX
XX Synthetic construct PR975YM SEQ ID NO:31.
XX
XX Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX
XX Human immunodeficiency virus type C.
OS Synthetic.
OS
PN WO200204493-A2.
XX
XX 17-JAN-2002.
XX
XX 05-JUL-2001; 2001WO-US21241.
XX
XX 05-JUL-2000; 2000US-0610313.
XX
XX (CHIR) CHIRON CORP.
PA HIV FS(-)-ProtMod
PA Modified HIV prote
PI HIV FS(-)-ProtMod
PI HIV FS(-)-ProtMod
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX WPI; 2002-154920/20.
XX

PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
XX cell lines, particularly in gene therapy -
XX
XX Claim 1; Fig 9; 233pp; English.
CC The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding gag, pol, vif, vpr, tat, rev, vpu, env or
CC nef (1). (i) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV Type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the
CC exemplification of the present invention.
XX
XX Sequence 2463 BP; 567 A; 835 C; 759 G; 302 T; 0 other;
SQ

Query Match 100.0%; Score 2463; DB 24; Length 2463;
Best Local Similarity 100.0%; Pred. No. 1.4e-300;
Matches 2463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCAGCCACCATGGCGGAGCCATGAGCGAGCCAGCCAGCCACATCCTGATGAG 60
DB 1 GTCAGCCACCATGGCGGAGCCATGAGCGAGCCAGCCAGCCACATCCTGATGAG 60
QY 61 GCGAGCACTTCAAGGGCCCAAGCGCATCATCAAGTGTTCACACTGGGCAAGAGGGC 120
DB 61 GCGAGCACTTCAAGGGCCCAAGCGCATCATCAAGTGTTCACACTGGGCAAGAGGGC 120
QY 121 CACATCGCCCACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAG 180
DB 121 CACATCGCCCACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAG 180
QY 181 GCGCACCAGATGAAGGACTGCACCGAGCGCGAGCGCCCAACTTCTTCGCGAGGAGCTGGCC 240
DB 181 GCGCACCAGATGAAGGACTGCACCGAGCGCGAGCGCCCAACTTCTTCGCGAGGAGCTGGCC 240
QY 241 TTCCCGCAGGCAAGGCGCGGAGTTCGCCAGCGAGGAGGAGCGCCGCAAGCGCCAC 300
DB 241 TTCCCGCAGGCAAGGCGCGGAGTTCGCCAGCGAGGAGGAGCGCCGCAAGCGCCAC 300
QY 301 AGCGCGAGCTGCAGGTGCGCGGGGCAACCCCGCAGCGAGCGCGCGCGAGCGCCAG 360
DB 301 AGCGCGAGCTGCAGGTGCGCGGGGCAACCCCGCAGCGAGCGCGCGAGCGCCAG 360
QY 361 GGCACCTTGAATTCCTCCCGAGATCACCCTGTGCGAGCGCCCGCTGGTGAATCAAGTG 420
DB 361 GGCACCTTGAATTCCTCCCGAGATCACCCTGTGCGAGCGCCCGCTGGTGAATCAAGTG 420
QY 421 GCGCGCCAGATCAAGGAGGCGCTGCTGACACCGCGCGCGAGCACACCGTGTGAGGAG 480
DB 421 GCGCGCCAGATCAAGGAGGCGCTGCTGACACCGCGCGCGAGCACACCGTGTGAGGAG 480
QY 481 ATGAGCCTTCCCGGCAAGTGAAGCCCAAGATGATGATGATGATGATGATGATGATGAT 540
DB 481 ATGAGCCTTCCCGGCAAGTGAAGCCCAAGATGATGATGATGATGATGATGATGATGAT 540
QY 541 GTGCGCCAGATGAGGAGGCGCTGCTGACACCGCGCGCGAGGAGGCGCCATGCGCACCGTG 600
DB 541 GTGCGCCAGATGAGGAGGCGCTGCTGACACCGCGCGAGGAGGCGCCATGCGCACCGTG 600
QY 601 CTGATCGCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCGAGCTGGGCTGC 660
DB 601 CTGATCGCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCGAGCTGGGCTGC 660
QY 661 ACCCTGAACTTCCCGCATCAGCGCCCATCGAGACCGCTGCGCGTGAAGCTGAAGCGCGGATG 720
DB 661 ACCCTGAACTTCCCGCATCAGCGCCCATCGAGACCGCTGCGCGTGAAGCTGAAGCGCGGATG 720
QY 721 GACGGCCCCAAGGTGAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCGCC 780

DB 721 GACGGCCCCAAGGTGAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCGCC 780
QY 781 ATCTGGAGGAGATGGAGAAGGAGGCAAGATCACCAGATCGGCCCGGAGAACCCCTAC 840
DB 781 ATCTGGAGGAGATGGAGAAGGAGGCAAGATCACCAGATCGGCCCGGAGAACCCCTAC 840
QY 841 AACACCCCTGTTCGCCATCAAGAAGAAGACAGCACCAGTGGCGCAAGCTGGTGGAC 900
DB 841 AACACCCCTGTTCGCCATCAAGAAGAAGACAGCACCAGTGGCGCAAGCTGGTGGAC 900
QY 901 TTCGCGAGGTGAACAAAGCGACCCAGACTTCTGGAGGTGACGTGGGATCCCGCCAC 960
DB 901 TTCGCGAGGTGAACAAAGCGACCCAGACTTCTGGAGGTGACGTGGGATCCCGCCAC 960
QY 961 CCGCGCGCTGGAAGAAGAAGAGCGTGACCTGTGACGTGGCGGAGCGCTACTTTC 1020
DB 961 CCGCGCGCTGGAAGAAGAAGAGCGTGACCTGTGACGTGGCGGAGCGCTACTTTC 1020
QY 1021 AGCTGCGCTTGGACGAGGACTTCCGCAAGTACACCGCTTTCACCATCCCGCAGCATCAAC 1080
DB 1021 AGCTGCGCTTGGACGAGGACTTCCGCAAGTACACCGCTTTCACCATCCCGCAGCATCAAC 1080
QY 1081 AACGAGACCCCGCGATCCGCTACAGTACACCTGACCTGCGCCAGGGCTGGAAGGGCAGC 1140
DB 1081 AACGAGACCCCGCGATCCGCTACAGTACACCTGACCTGCGCCAGGGCTGGAAGGGCAGC 1140
QY 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCGCCCAACCCC 1200
DB 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCGCCCAACCCC 1200
QY 1201 GAGATCGTGTATCTACAGGCCCCCTGTACGTGGGAGCGAGCTGGAGATCGGCCAGCAGC 1260
DB 1201 GAGATCGTGTATCTACAGGCCCCCTGTGTGTGGGAGCGAGCTGGAGATCGGCCAGCAGC 1260
QY 1261 CGCGCAGATCGAGGAGCTGCGCAAGCAGCTGTGCTGGGCTTCCACACCCCGCAGC 1320
DB 1261 CGCGCAGATCGAGGAGCTGCGCAAGCAGCTGTGCTGGGCTTCCACACCCCGCAGC 1320
QY 1321 AAGAAGCACCAGAGGAGCCCCCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
DB 1321 AAGAAGCACCAGAGGAGCCCCCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
QY 1381 TGGACCTGTGAGCCCATCGAGCTGCGCGAGAGGAGCTGGACCTGGAACGACATCCAG 1440
DB 1381 TGGACCTGTGAGCCCATCGAGCTGCGCGAGAGGAGCTGGACCTGGAACGACATCCAG 1440
QY 1441 AAGTGTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGGATCAAGGTGGCCAG 1500
DB 1441 AAGTGTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGGATCAAGGTGGCCAG 1500
QY 1501 CTGTGCAAGCTGTGCGCGGCGCCAAAGCCCTGACCGACATCGTCCCTGACCGAGGAG 1560
DB 1501 CTGTGCAAGCTGTGCGCGGCGCCAAAGCCCTGACCGACATCGTCCCTGACCGAGGAG 1560
QY 1561 GCGGAGCTGGAGCTGCGCGGAGAACCGCGAGATCTTGTGCGGAGCCGCTGACCGCTGTAC 1620
DB 1561 GCGGAGCTGGAGCTGCGCGGAGAACCGCGAGATCTTGTGCGGAGCCGCTGACCGCTGTAC 1620
QY 1621 TACGACCCCGCAGCAAGGAGCTGGTGGCGGAGATCCAGAAGCAGGCGCCAGACAGTGGAC 1680
DB 1621 TACGACCCCGCAGCAAGGAGCTGGTGGCGGAGATCCAGAAGCAGGCGCCAGACAGTGGAC 1680
QY 1681 TACGAGATCTACGAGGCGCTTCAAGACCTTGAAGACCGGCAAGTACCGCAAGTGGCG 1740
DB 1681 TACGAGATCTACGAGGCGCTTCAAGACCTTGAAGACCGGCAAGTACCGCAAGTGGCG 1740
QY 1741 ACCGCGCCACACCAAGCAGCTGAAGCAGCTGACGAGGCGCTGACAGAGTCCGCTTGGAG 1800
DB 1741 ACCGCGCCACACCAAGCAGCTGAAGCAGCTGACGAGGCGCTGACAGAGTCCGCTTGGAG 1800
QY 1801 AGCATCGTATCTGGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAGGAGACCTGG 1860

Db 1801 AGCATCGTGATCTGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAAAGGAGACCTGG 1860
Qy 1861 GAGACCTGGTGGACCGACTACTGGCAGCCACCTGTGATCCCGAGTGGAGTTCTGTGAC 1920
Db 1861 GAGACCTGGTGGACCGACTACTGGCAGCCACCTGTGATCCCGAGTGGAGTTCTGTGAC 1920
Qy 1921 ACCCCCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGCCCGAG 1980
Db 1921 ACCCCCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGCCCGAG 1980
Qy 1981 ACCTTCTACGTGGAGCGCGCCCAACCGGAGACCAAGATCGGCAAGCGCGGTACGTG 2040
Db 1981 ACCTTCTACGTGGAGCGCGCCCAACCGGAGACCAAGATCGGCAAGCGCGGTACGTG 2040
Qy 2041 ACCGACCGGGCGGCGAGAGATCGTAGCTGACCGAGACCAACCAAGAACCGAG 2100
Db 2041 ACCGACCGGGCGGCGAGAGATCGTAGCTGACCGAGACCAACCAAGAACCGAG 2100
Qy 2101 CTGACGGCCATCCAGCTGGCCCTCGAGGACAGCGGCGAGGTTGAACATCGTGACCGAC 2160
Db 2101 CTGACGGCCATCCAGCTGGCCCTCGAGGACAGCGGCGAGGTTGAACATCGTGACCGAC 2160
Qy 2161 AGCCAGTACGCGCTGGGCATCATCAGGCCCGAGCGGAGGAGCGAGCGTGGTG 2220
Db 2161 AGCCAGTACGCGCTGGGCATCATCAGGCCCGAGCGGAGGAGCGAGCGTGGTG 2220
Qy 2221 AACCATGATCATCGACGAGCTGATCAAGAGGAGAGGTTACCTGAGCTGGTGCCTGCC 2280
Db 2221 AACCATGATCATCGACGAGCTGATCAAGAGGAGAGGTTACCTGAGCTGGTGCCTGCC 2280
Qy 2281 CACAAGGCGATCGGCGGCAAGCAGAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2340
Db 2281 CACAAGGCGATCGGCGGCAAGCAGAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAG 2340
Qy 2341 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAAGTACATGAGACACCTG 2400
Db 2341 GTGCTGTTCCTGGACGGCATCGATGGCGGCATCGTGATCTACCAAGTACATGAGACACCTG 2400
Qy 2401 TACGTGGGACGGCGGCGCTTAGGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAA 2460
Db 2401 TACGTGGGACGGCGGCGCTTAGGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAA 2460
Qy 2461 TTC 2463
Db 2461 TTC 2463

RESULT 2

ABL39959
ID ABL39959 standard; DNA; 2469 BP.
XX ABL39959;
AC
XX
XX
DT 15-MAY-2002 (first entry)
XX
DE Synthetic construct PR975(+) SEQ ID NO:30.
XX
XX Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX
XX Human immunodeficiency virus type C.
OS Synthetic.
XX
XX WO200204493-A2.
XX
XX 17-JAN-2002.
PD
XX
XX 05-JUL-2001; 2001WO-US21241.
PF
XX
XX 05-JUL-2000; 2000US-0610313.
PR
XX
XX (CHIR) CHIRON CORP.

PA

(UYST-) UNIV STELLENBOSCH.

Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;

WPI; 2002-154920/20.

New polynucleotides encoding antigenic HIV Type C polypeptides, useful in applications including DNA immunization or generation of packaging cell lines, particularly in gene therapy

Claim 1; Fig 8; 233pp; English.

The present invention describes expression cassettes comprising a polynucleotide sequence encoding a polypeptide comprising immunogenic HIV type C polypeptides. The expression cassettes comprise any of the HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef. (i) have immunostimulant activity and can be used in gene therapy. The HIV type C polynucleotides are useful in applications including DNA immunisation, generation of packaging cell lines, and production of HIV Type C proteins. The polynucleotides are particularly useful in gene therapy and DNA immunisation applications. ABL39942 to ABL40054 and ABL406204 to ABL406215 represent sequences used in the exemplification of the present invention.

Sequence 2469 BP; 571 A; 833 C; 761 G; 304 T; 0 other;

Query Match 99.2%; Score 2442.2; DB 24; Length 2469;

Best Local Similarity 99.6%; Pred. No. 5.6e-298;

Matches 2460; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

Qy 1 GTGACGCGCCACCATGCGGAGGCCATGAGCCAGGCCACCGCCCAACATCTCTGATGCAG 60

Db 1 GTGACGCGCCACCATGCGGAGGCCATGAGCCAGGCCACCGCCCAACATCTCTGATGCAG 60

Qy 61 GCGAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGC 120

Db 61 GCGAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGC 120

Qy 121 CACATCGCCCGCAACTGCGGCGCCCCCGCAGAGAGGCTGCTGAAAGTGGCGCAAGGAG 180

Db 121 CACATCGCCCGCAACTGCGGCGCCCCCGCAGAGAGGCTGCTGAAAGTGGCGCAAGGAG 180

Qy 181 GCGCACCAGATGAAGGACTGCAACCGAGCGGCCAGGCAACTTCTCCGCGAGGACCTGGCC 240

Db 181 GCGCACCAGATGAAGGACTGCAACCGAGCGGCCAGGCAACTTCTCCGCGAGGACCTGGCC 240

Qy 241 TTCCCCAGGGAGCGCCCGGAGTTCCCGAGCGAGAGACCGCGCCCAACAGCCCGACCC 300

Db 241 TTCCCCAGGGAGCGCCCGGAGTTCCCGAGCGAGAGACCGCGCCCAACAGCCCGACCC 300

Qy 301 AGCGCGAGCTGCGAGTGGCGGCGACAAACCCCGCAGGAGCGCGCGCGCGCGCCAG 360

Db 301 AGCGCGAGCTGCGAGTGGCGGCGACAAACCCCGCAGGAGCGCGCGCGCGCGCCAG 360

Qy 361 GGCACCTTGAATTCCTCCCGAGATCACCTGTGGCAGCGCCCTGCTGAGCATCAAGGTG 420

Db 361 GGCACCTTGAATTCCTCCCGAGATCACCTGTGGCAGCGCCCTGCTGAGCATCAAGGTG 420

Qy 421 GCGGCGCAGATCAAGGAGGCGCTGCTGACACCGCGCGCGAGACACCGTGTGGAGGAG 480

Db 421 GCGGCGCAGATCAAGGAGGCGCTGCTGACACCGCGCGCGAGACACCGTGTGGAGGAG 480

Qy 481 ATGAGCTTGGCGCGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGGCTTTCATCAAG 540

Db 481 ATGAGCTTGGCGCGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGGCTTTCATCAAG 540

Qy 541 GTGGCGCAGTACGACCCAGATCTGATCGAGATCTGCGGCAAGAGCCCATCGGACCGCTG 600

Db 541 GTGGCGCAGTACGACCCAGATCTGATCGAGATCTGCGGCAAGAGCCCATCGGACCGCTG 600

Qy 601 CTGATCGGCGCGCCCGCTGAAACATCATCGGCGCGCAACATGCTGACCCAGTGGGCTGC 660

Db 601 CTGATCGGCGCGCCCGCTGAAACATCATCGGCGCGCAACATGCTGACCCAGTGGGCTGC 660


```
QY 661 ACCTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGAAGCTGAAGCCCGCGCATG 720
Db 661 ACCTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGAAGCTGAAGCCCGCGCATG 720
QY 721 GACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAGATCAAGGGCCCTGACCGCC 780
Db 721 GACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAGATCAAGGGCCCTGACCGCC 780
QY 781 ATCTGGAGAGATGGAGAGAGGGCAAGATCACCAGATCGGCCCGCCGAGAACCCCTAC 840
Db 781 ATCTGGAGAGATGGAGAGAGGGCAAGATCACCAGATCGGCCCGCCGAGAACCCCTAC 840
QY 841 AACACCCCTGTTGCGCATCAAGAAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
Db 841 AACACCCCTGTTGCGCATCAAGAAAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
QY 901 TTCGGAGAGTGAACACAGCCACCCAGGACTTCTGGAGAGTGCAGCTGGGCATCCCCAC 960
Db 901 TTCGGAGAGTGAACACAGCCACCCAGGACTTCTGGAGAGTGCAGCTGGGCATCCCCAC 960
QY 961 CCGCGGCGCTGAGAGAGAGAGAGCGTGACCGTCTGACAGTGGCGAGCGCTACTTC 1020
Db 961 CCGCGGCGCTGAGAGAGAGAGAGCGTGACCGTCTGACAGTGGCGAGCGCTACTTC 1020
QY 1021 AGCGTGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTACCATCCCGCAGCATCAAC 1080
Db 1021 AGCGTGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTACCATCCCGCAGCATCAAC 1080
QY 1081 AACGAGACCCCGCGCATCCGCTACCAAGTACACCGTGTGCGCCAGGGCTGGAAGGGCAGC 1140
Db 1081 AACGAGACCCCGCGCATCCGCTACCAAGTACACCGTGTGCGCCAGGGCTGGAAGGGCAGC 1140
QY 1141 CCCAGCATCTTCCAGACGACATGACCAAGATCCTGAGAGCCCTTCCGCGCCGCAACCCC 1200
Db 1141 CCCAGCATCTTCCAGACGACATGACCAAGATCCTGAGAGCCCTTCCGCGCCGCAACCCC 1200
QY 1201 GAGATCGTGATCTACCA-----GGCCCCCTGTACCTGGCGACGACCTGGAGATCGGC 1254
Db 1201 GAGATCGTGATCTACCAAGTACATGGACGACCTGTACCTGGCGACGACCTGGAGATCGGC 1254
QY 1255 CAGCACCGCCGCAAGATCGAGGAGCTGCGCAAGACCTGTGGCTGGGGCTTACCAACC 1314
Db 1255 CAGCACCGCCGCAAGATCGAGGAGCTGCGCAAGACCTGTGGCTGGGGCTTACCAACC 1314
QY 1315 CCGACAAGAACACCAAGAGAGCCCTTCTGCTGGATGGCTACGAGCTGCACCCC 1374
Db 1315 CCGACAAGAACACCAAGAGAGCCCTTCTGCTGGATGGCTACGAGCTGCACCCC 1374
QY 1375 GACAAGTGGACCGCTGACCCCATCGAGCTGCCGAGAGAGAGCTGGACCCGTGAACGAC 1434
Db 1375 GACAAGTGGACCGCTGACCCCATCGAGCTGCCGAGAGAGAGCTGGACCCGTGAACGAC 1434
QY 1435 ATCCAGAAGCTGTGGGCAAGCTGAAGTGGCGCAGCCAGATCTACCCCGGCATCAAGGTG 1494
Db 1435 ATCCAGAAGCTGTGGGCAAGCTGAAGTGGCGCAGCCAGATCTACCCCGGCATCAAGGTG 1494
QY 1495 CGCCAGCTGTGCAAGCTGTGCGCGGCGCAAGCCCTGACCGACATCTGTCGCCCTGACC 1554
Db 1495 CGCCAGCTGTGCAAGCTGTGCGCGGCGCAAGCCCTGACCGACATCTGTCGCCCTGACC 1554
QY 1555 GAGAGCCGAGCTGGAGCTGGCGGAGAACCGGAGATCTTCCGCGAGCCCGCTGACCGGC 1614
Db 1555 GAGAGCCGAGCTGGAGCTGGCGGAGAACCGGAGATCTTCCGCGAGCCCGCTGACCGGC 1614
QY 1615 GTGTACTACCAACCCACAGGACCTGGTGGCGGAGATCAGAAAGAGAGGCGCAGCAG 1674
Db 1615 GTGTACTACCAACCCACAGGACCTGGTGGCGGAGATCAGAAAGAGAGGCGCAGCAG 1674
QY 1675 TGGACCTACCAAGTCTACCAAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGCCAAG 1734
Db 1675 TGGACCTACCAAGTCTACCAAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGCCAAG 1734
QY 1735 ATGCGACACCCACACCAACAGAGCTGAAGCAGCTGACCGAGGCGCTGCAAGAGATCGCC 1794
Db 1735 ATGCGACACCCACACCAACAGAGCTGAAGCAGCTGACCGAGGCGCTGCAAGAGATCGCC 1794
QY 1795 ATGAGAGACATCGTGTCTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAGAG 1854
Db 1795 ATGAGAGACATCGTGTCTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAGAG 1854
QY 1855 ACCTGGAGACCTGGTGGACCGACTACTGCGAGCCACCTGGATCCCGAGTGGAGTTC 1914
Db 1855 ACCTGGAGACCTGGTGGACCGACTACTGCGAGCCACCTGGATCCCGAGTGGAGTTC 1914
QY 1915 GTAAACACCCCCCTGGTGAAGCTGTGTGTTACAGCTGGAGAGAGCCCATCATCGGC 1974
Db 1915 GTAAACACCCCCCTGGTGAAGCTGTGTGTTACAGCTGGAGAGAGCCCATCATCGGC 1974
QY 1975 GCGAGACCTTCTACGTGGAGCGGCCGCCAACCGCAGACCAAGATCGGCAAGCGCGGC 2034
Db 1975 GCGAGACCTTCTACGTGGAGCGGCCGCCAACCGCAGACCAAGATCGGCAAGCGCGGC 2034
QY 2035 TAGCTGACCCAGCGGCGCGCAGAAAGATCGTGAAGCTGACCGAGACCAACCAAGAG 2094
Db 2035 TAGCTGACCCAGCGGCGCGCAGAAAGATCGTGAAGCTGACCGAGACCAACCAAGAG 2094
QY 2095 ACCGAGCTGACGAGCCATCCAGCTGGCCCTGCAGGACAGCGCAGAGGTGAACATCGTG 2154
Db 2095 ACCGAGCTGACGAGCCATCCAGCTGGCCCTGCAGGACAGCGCAGAGGTGAACATCGTG 2154
QY 2155 ACCGACAGCAGTACGCGCTGGGCTATCCAGGCCCAGCCCGCAGAGCGAGCGAG 2214
Db 2155 ACCGACAGCAGTACGCGCTGGGCTATCCAGGCCCAGCCCGCAGAGCGAGCGAG 2214
QY 2215 CTGTTGAACACAGATCATCGAGCAGCTCATCAAGAAAGAGAGGTGTACCTGAGCTGGGTG 2274
Db 2215 CTGTTGAACACAGATCATCGAGCAGCTCATCAAGAAAGAGAGGTGTACCTGAGCTGGGTG 2274
QY 2275 CCGCCCCACAAGGSCATCGCGGCAACGACAGCATGACAAAGCTGGTGAAGAGGGCATC 2334
Db 2275 CCGCCCCACAAGGSCATCGCGGCAACGACAGCATGACAAAGCTGGTGAAGAGGGCATC 2334
QY 2335 CGCAAGTGTGTCTTCTGGAGCGCATCGATGGCGCATCGTGTATCTACCAAGTACATGGAC 2394
Db 2335 CGCAAGTGTGTCTTCTGGAGCGCATCGATGGCGCATCGTGTATCTACCAAGTACATGGAC 2394
QY 2395 GACCTGTACCTGGCGAGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACC 2454
Db 2395 GACCTGTACCTGGCGAGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACC 2454
QY 2455 GGTGAATTC 2463
Db 2455 GGTGAATTC 2469
```

RESULT 3

ABL39961

ID ABL39961 standard; DNA; 2457 BP.

XX

XX

XX

XX

DT 15-MAY-2002 (first entry)

XX

DE Synthetic construct PR975YMMW SEQ ID NO:32.

XX

KW Human immunodeficiency virus type C; antigenic HIV type C protein;

KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; env; nef;

KW immunostimulant; gene therapy; gene; ds.

XX

OS Human immunodeficiency virus type C.

OS Synthetic.

XX

PN WO200204493-A2.

XX

PD 17-JAN-2002.

XX 05-JUL-2001; 2001WO-US21241.
XX PF
XX PR
XX XX
XX (CHIR) CHIRON CORP.
XX PA (UYST-) UNIV STELLENBOSCH.
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX WPI; 2002-154920/20.
XX
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful
XX PT in applications including DNA immunization or generation of packaging
XX PT cell lines, particularly in gene therapy -
XX PS
XX Claim 1; Fig 10; 233pp; English.
XX
XX The present invention describes expression cassettes comprising a
XX polynucleotide sequence encoding a polypeptide comprising immunogenic
XX HIV type C polypeptides. The expression cassettes comprise any of the
XX HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
XX Nef (1). (1) have immunostimulant activity and can be used in gene
XX therapy. The HIV type C polynucleotides are useful in applications
XX including DNA immunisation, generation of packaging cell lines, and
XX production of HIV type C proteins. The polynucleotides are particularly
XX useful in gene therapy and DNA immunisation applications. ABL39942 to
XX ABL40054 and ABL06204 to ABL06215 represent sequences used in the
XX exemplification of the present invention.
XX SQ Sequence 2457 BP; 566 A; 837 C; 754 G; 300 T; 0 other;

Query Match 98.9%; Score 2436.2; DB 24; Length 2457;
Best Local Similarity 99.6%; Pred. No. 3.2e-257;
Matches 2454; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 1 GTCAGCGCCACCATGGCGAGGCCATGAGCGAGGCCACCGAGCGCCCAACATCCTGATGCAG 60
DB 1 GTCAGCGCCACCATGGCGAGGCCATGAGCGAGGCCACCGAGCGCCCAACATCCTGATGCAG 60

QY 61 CGCAGCAACTTCAAGGCGCCCAAGCGCATCATCAAGTCTCAACTGCGGCAAGGAGGC 120
DB 61 CGCAGCAACTTCAAGGCGCCCAAGCGCATCATCAAGTCTCAACTGCGGCAAGGAGGC 120

QY 121 CACATCGCCGCAACTGCCGCGCCCGCCGCAAGAGGCGTCTGGAAGTGGCGCAAGGAG 180
DB 121 CACATCGCCGCAACTGCCGCGCCCGCCGCAAGAGGCGTCTGGAAGTGGCGCAAGGAG 180

QY 181 GGCACACAGATGAAGGACTGACCGAGCGCCAGCGCAACTTCTTCGCGAGGACCTGGCC 240
DB 181 GGCACACAGATGAAGGACTGACCGAGCGCCAGCGCAACTTCTTCGCGAGGACCTGGCC 240

QY 241 TTCCCGCAGGCAAGCGCCGAGTTCCTCCAGCGAGCAGACCGCGCCCAAGCGCCACC 300
DB 241 TTCCCGCAGGCAAGCGCCGAGTTCCTCCAGCGAGCAGACCGCGCCCAAGCGCCACC 300

QY 301 AGCGCGAGGTGCAAGTGCAGCGGCGCAACCCCGCAGCGAGCGCGCGCGCGCGCCAG 360
DB 301 AGCGCGAGGTGCAAGTGCAGCGGCGCAACCCCGCAGCGAGCGCGCGCGCGCGCCAG 360

QY 361 GGCACCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCCCTTGGTGAAGATCAAGTG 420
DB 361 GGCACCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCCCTTGGTGAAGATCAAGTG 420

QY 421 GGCAGCGAGATCAAGGAGGCGCTCTGGACACCGCGCGCGCGCGCGCGCGCGCGAG 480
DB 421 GGCAGCGAGATCAAGGAGGCGCTCTGGACACCGCGCGCGCGCGCGCGCGCGAG 480

QY 481 ATGAGCTGCGCGCAAGTGGAGCCCAAGATGATCGCGCGCATCGCGGCTTTCATCAAG 540
DB 481 ATGAGCTGCGCGCAAGTGGAGCCCAAGATGATCGCGCGCATCGCGGCTTTCATCAAG 540

QY 541 GTGCGCCAGTACGACCAAGATCTCTGATCGAGATCTGCGGCAAGAGGCGCATCGGCAACGCTG 600

DB 541 GTGCGCCAGTACGACCAAGATCTCTGATCGAGATCTCGCGCAAGAGGCGCATCGGCAACGCTG 600
QY 601 CTGATCGGCGCCACCCCGTGAACATCATCGGCGCCACACATGCTACCCAGCTGGGCTGC 660
DB 601 CTGATCGGCGCCACCCCGTGAACATCATCGGCGCCACACATGCTACCCAGCTGGGCTGC 660
QY 661 ACCCTGAATTTCCCATCAGCCCATCGAGACCGTGCCTGAACTGAAGCTGAAGCCCGGCATG 720
DB 661 ACCCTGAATTTCCCATCAGCCCATCGAGACCGTGCCTGAACTGAAGCTGAAGCCCGGCATG 720
QY 721 GACGCCCCCAAGGTGAAGCAGTGGCCCTCGACCGAGGAGAGATCAAGGCCCTTGACCCGCC 780
DB 721 GACGCCCCCAAGGTGAAGCAGTGGCCCTCGACCGAGGAGAGATCAAGGCCCTTGACCCGCC 780
QY 781 ATCTCGAGGAGATGGAGAAGAGGAGGCAAGATCACCAGATCGGCCCGCGAGAACCCCTAC 840
DB 781 ATCTCGAGGAGATGGAGAAGAGGAGGCAAGATCACCAGATCGGCCCGCGAGAACCCCTAC 840
QY 841 AACACCCCGCTGTTCCGCCATCAAGAAAGAGACACCAAGTGGCGCAAGCTGGTGGAC 900
DB 841 AACACCCCGCTGTTCCGCCATCAAGAAAGAGACACCAAGTGGCGCAAGCTGGTGGAC 900
QY 901 TTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGATCCGCCAC 960
DB 901 TTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGATCCGCCAC 960
QY 961 CCGCGCGCTTGAAGAAGAAAGAGCGTGCCTGCTGGAGCTGGGCGAGCGCTTACTTC 1020
DB 961 CCGCGCGCTTGAAGAAGAAAGAGCGTGCCTGCTGGAGCTGGGCGAGCGCTTACTTC 1020
QY 1021 AGCTGTCCTTGGAGGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGAGCATCAAC 1080
DB 1021 AGCTGTCCTTGGAGGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGAGCATCAAC 1080
QY 1081 AACGAGACCCCGCGCATCCGCTACCAAGTACAACTGCTGCCCGAGGCTGGAAGGCGAGC 1140
DB 1081 AACGAGACCCCGCGCATCCGCTACCAAGTACAACTGCTGCCCGAGGCTGGAAGGCGAGC 1140
QY 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCGCAACGCC 1200
DB 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCGCAACGCC 1200
QY 1201 GAGATCGTGATCTACCAAGCGCCCGCTGTACGTGGGACGAGCTGGAGATCGGCGAGCAG 1260
DB 1201 GAGATCGTGATCTACCAAGCGCCCGCTGTACGTGGGACGAGCTGGAGATCGGCGAGCAG 1260
QY 1261 CGCGCAAGATCGAGGAGCTGCGCAAGCACCCTGTGCGCTGGGGCTTCCACACCCCGCAG 1320
DB 1261 CGCGCAAGATCGAGGAGCTGCGCAAGCACCCTGTGCGCTGGGGCTTCCACACCCCGCAG 1320
QY 1321 AAGAAGCACAGAGGAGGCGCCCTTCTGTGGATGGGCTAGGAGCTGCACCCCGCAAG 1380
DB 1321 AAGAAGCACAGAGGAGGCGCCCTTCTGTGGATGGGCTAGGAGCTGCACCCCGCAAG 1380
QY 1381 TGGACCGTGCAGCGCATCGAGCTGCCGAGAGGAGAGCTGACCGTGAACGACATCCAG 1440
DB 1381 TGGACCGTGCAGCGCATCGAGCTGCCGAGAGGAGAGCTGACCGTGAACGACATCCAG 1440
QY 1441 AAGCTGTGGCAAGCTGAACCTGGGCGCCAGCGAGATCTACCCCGGATCAAGGTGCGCCAG 1500
DB 1441 AAGCTGTGGCAAGCTGAACCTGGGCGCCAGCGAGATCTACCCCGGATCAAGGTGCGCCAG 1500
QY 1501 CTGTGTGCAAGCTGTGCGCGCGCGCAAGGCCCTGACCGACATCTGTGCCCTGACCGAGGAG 1560
DB 1501 CTGTGTGCAAGCTGTGCGCGCGCGCAAGGCCCTGACCGACATCTGTGCCCTGACCGAGGAG 1560
QY 1561 GCCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTCGCGAGCGCGCTGCAGCGGCTGTAC 1620
DB 1561 GCCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTCGCGAGCGCGCTGCAGCGGCTGTAC 1620
QY 1621 TACGACCCCGCAGAGGAGCTGTGTGCGCGCGAGATCCAGAGAGCGGCGCAGGACAGTGGACC 1680
DB 1621 TACGACCCCGCAGAGGAGCTGTGTGCGCGCGAGATCCAGAGAGCGGCGCAGGACAGTGGACC 1680

Db 1615 TACGACCCAGCAGGACCTGTGTGCGGAGATCCAGAAGCAGGGCCACGACCAAGTGGACC 1674
QY 1681 TACCAGATCTACAGAGAGCCCTTCAGAAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1740
Db 1675 TACCAGATCTACAGAGAGCCCTTCAGAAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1734
QY 1741 ACCGCCACACCAACACGCTGAAGCAGCTGACCGAGGCGGTGAGAAAGATCGCATGGAG 1800
Db 1735 ACCGCCACACCAACACGCTGAAGCAGCTGACCGAGGCGGTGAGAAAGATCGCATGGAG 1794
QY 1801 AGCATCGTGATCTGGGGCAAGACCCCAAGTTCCTGGCTGCCCATCCAGAAAGACCTGG 1860
Db 1795 AGCATCGTGATCTGGGGCAAGACCCCAAGTTCCTGGCTGCCCATCCAGAAAGACCTGG 1854
QY 1861 GAGACCTGGTGACCACTACTCTGGCAGGCCACCTGGATCCCCAGTGGAGTTCGTGAAC 1920
Db 1855 GAGACCTGGTGACCACTACTCTGGCAGGCCACCTGGATCCCCAGTGGAGTTCGTGAAC 1914
QY 1921 ACCCCCCCTCTGTGAAGCTGTGTACCACTGTGGAAGAGGCCCATCATCGGCGCGGAG 1980
Db 1915 ACCCCCCCTCTGTGAAGCTGTGTACCACTGTGGAAGAGGCCCATCATCGGCGCGGAG 1974
QY 1981 ACCTTTACGTGACGGCGCGCCCAACCGCGAGACCAAGATCGGCAGAGCGCGCTACGTG 2040
Db 1975 ACCTTTACGTGACGGCGCGCCCAACCGCGAGACCAAGATCGGCAGAGCGCGCTACGTG 2034
QY 2041 ACCGACCGGGCGCGCAGAGATCTGTGACCTGACCGAGACCAACCAAGAGACCGAG 2100
Db 2035 ACCGACCGGGCGCGCAGAGATCTGTGAGCTGACCGAGACCAACCAAGAGACCGAG 2094
QY 2101 CTGACGGCCATCCAGCTGGCCCTGCGAGGACAGCGGAGAGTGGAACATCGTGACCGAC 2160
Db 2095 CTGACGGCCATCCAGCTGGCCCTGCGAGGACAGCGGAGAGTGGAACATCGTGACCGAC 2154
QY 2161 AGCCAGTACGCCCTGGGCATCATCAGGCCCGCCGACAAAGAGCAGAGCGAGCTGGTG 2220
Db 2155 AGCCAGTACGCCCTGGGCATCATCAGGCCCGCCGACAAAGAGCAGAGCGAGCTGGTG 2214
QY 2221 AACCATGATCATCAGCAGCTGTGATCAAGAGGAGAGTGTTACCTGAGCTGGTGCCCGCC 2280
Db 2215 AACCATGATCATCAGCAGCTGTGATCAAGAGGAGAGTGTTACCTGAGCTGGTGCCCGCC 2274
QY 2281 CACAAGGGCATCGCGCGGCAACGACGAGATCGACAAGCTGGTGAGCAAGGSCATCCGCAAG 2340
Db 2275 CACAAGGGCATCGCGCGGCAACGACGAGATCGACAAGCTGGTGAGCAAGGSCATCCGCAAG 2334
QY 2341 GTGCTGTTCTGACCGCATCGATGCGGCGCATCGTGATCTACCACTACATGGAGACCTG 2400
Db 2335 GTGCTGTTCTGACCGCATCGATGCGGCGCATCGTGATCTACCACTACATGGAGACCTG 2394
QY 2401 TAGCTGGGACGGCGGCCCTAGGATCGATTAAAGCTTCCCGGGGCTAGCACCGGTGAA 2460
Db 2395 TAGCTGGGACGGCGGCCCTAGGATCGATTAAAGCTTCCCGGGGCTAGCACCGGTGAA 2454
QY 2461 TTC 2463
Db 2455 TTC 2457

RESULT 4
ID AAA70479
XX AAA70479 standard; DNA; 2306 BP.
AC AAA70479;
XX
XX
XX 28-NOV-2000 (first entry)
XX
XX HIV FS(-)_ProtMod_RTOpt_YM coding sequence.
XX HIV-1; AIDS; Gag-reverse transcriptase; vaccine; expression cassette; ss.
XX Human immunodeficiency virus type 1.
OS Synthetic.

XX WO200039302-A2.
PN 06-JUL-2000.
XX 30-DEC-1999; 99WO-US31245.
XX 31-DEC-1998; 98US-0114495.
PR 01-DEC-1999; 99US-0168471.
XX (CHIR) CHIRON CORP.
XX Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;
PI Greer C, Selby M, Walker C;
XX WPI; 2000-452400/39.
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
PT Gag-containing polypeptide useful for vaccinating against HIV
PT infections and acquired immunodeficiency syndrome (AIDS) -
XX Claim 7; Fig 73; 391pp; English.
XX The present sequence is the coding sequence of a synthetic HIV
CC Gag-reverse transcriptase expression cassette, FS(-)_ProtMod_RTOpt_YM.
CC The Gag protein of HIV is needed for the assembly of virus-like
CC particles. In addition, the Gag protein is involved in many stages of the
CC HIV life cycle, including assembly, virion maturation after particle
CC release and early post-entry steps in viral replication. The expression
CC cassette may be used for the recombinant expression of HIV
CC Gag-polypeptides which may then be used to vaccinate against HIV
CC infection and acquired immunodeficiency syndrome (AIDS).
XX Sequence 2306 BP; 529 A; 752 C; 730 G; 295 T; 0 other;
SQ
Query Match 83.1%; Score 2046; DB 21; Length 2306;
Best Local Similarity 93.6%; Pred. No. 2.4e-248;
Matches 2159; Conservative 0; Mismatches 135; Indels 12; Gaps 2;
QY 170 GCGGCAAGGAGGCCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAAGTCTTCGCGC 229
Db 1 GCGGCGCGGAGGACACCAAAATGAAAGATTGCATGAGAGACAGGCTAATTCTTCGCGC 60
QY 230 AGGACCTGGCTTCCCGCAGGCAAGCGCGAGTTCGCCAGCAGCAGCAAGCGGCCA 289
Db 61 AGAAGCTGGCTTCTCCAGGCAAGCGCGCGAGTTTCAGCAGCAGCAGCAGCGGCCA 120
QY 290 ACAGCCCCACAGCGCGAGCTGCAGTGCAGCGCGGCGGCGGCGGCGGCGGCGG 343
Db 121 ACAGCCCCACCGCGCGAGCTGCAGTGCAGCGCGGCGGCGGCGGCGGCGGCGG 180
QY 344 CCGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 397
Db 181 CCGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
QY 398 GCGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 457
Db 241 GCGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
QY 458 CCGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 517
Db 301 CCGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360
QY 518 GCGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 577
Db 361 GCGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 578 GCAAGAGGCGCATCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 637
Db 421 GCGGCGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 638 ACATGCTGACCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 697
XX

Db 481 ACCTGCTGACCCAGATCGGCTGCACCTTGAACCTTCCCATCAGCCCCATCGAGACGGTGC 540
QY 698 CCCTGAAGCTGAAGCCCGGATGACGCGCCCAAGGTGAAGCACTGCCCCGTGACCGAGG 757
Db 541 CCCTGAAGCTGAAGCCCGGATGACGCGCCCAAGGTGAAGCACTGCCCCGTGACCGAGG 600
QY 758 AGAAGATCAAGGCCCTGACCCGCTGCGAGGAGATGGAGAAGGAGGCGCAAGATCACCA 817
Db 601 AGAAGATCAAGGCCCTGCTGGAGATCTGCACCGAGATGGAGAAGGAGGCGCAAGATCACCA 660
QY 818 AGATCGGCCCGGAGAACCCCTACAAACCCCGCTGTTGCGCATCAAGAAGAGACAGCA 877
Db 661 AGATCGGCCCGGAGAACCCCTACAAACCCCGCTGTTGCGCATCAAGAAGAGACAGCA 720
QY 878 CCAAGTGGCGCAAGCTGGTGACTTCCGAGCTGAACAAAGCGACCCAGGACTTCTGGG 937
Db 721 CCAAGTGGCGCAAGCTGGTGACTTCCGAGCTGAACAAAGCGACCCAGGACTTCTGGG 780
QY 938 AGGTGACGTGGGCATCCCCACCCCGCGCTTGAAGAAGAGAGCGTGACCCGTGC 997
Db 781 AGGTGACGTGGGCATCCCCACCCCGCGCTTGAAGAAGAGAGCGTGACCCGTGC 840
QY 998 TGGAGCTGGCGGCGCTTACTTACAGCTGCCCTGGAGGAGCTTCCGCAAGTACACCG 1057
Db 841 TGGAGCTGGCGGCGCTTACTTACAGCTGCCCTGGAGGAGCTTCCGCAAGTACACCG 900
QY 1058 CTTTACCATCCCCAGCATCAACAAAGAGACCCCGCGCTACCAAGTACAACTGC 1117
Db 901 CTTTACCATCCCCAGCATCAACAAAGAGACCCCGCGCTACCAAGTACAACTGC 960
QY 1118 TGCCCCAGGGGTGAAGGGGAGCCCCAGCATCTTCCAGAGCAGATGACCAAGATCCTGG 1177
Db 961 TGCCCCAGGGGTGAAGGGGAGCCCCAGCATCTTCCAGAGCAGATGACCAAGATCCTGG 1020
QY 1178 AGCCCTTCCGCGCCGCAACCCAGATCTGTGATCTACCAAGGCCCCCTGTAGCTGGGCA 1237
Db 1021 AGCCCTTCCGCGCCGCAACCCAGATCTGTGATCTACCAAGGCCCCCTGTAGCTGGGCA 1080
QY 1238 GCGACCTGGAGATCGGCGAGCAGCCGCGCAAGATCGAGGAGCTCGCAAGCACTGCTGC 1297
Db 1081 GCGACCTGGAGATCGGCGAGCAGCCGCGCAAGATCGAGGAGCTCGCAAGCACTGCTGC 1140
QY 1298 GCTGGGGCTTCAACACCCCGCAAGAAGCAACCAAGAGGAGGCCCCCTTCTGTGATGG 1357
Db 1141 GCTGGGGCTTCAACACCCCGCAAGAAGCAACCAAGAGGAGGCCCCCTTCTGTGATGG 1200
QY 1358 GCTACGAGCTGCACCCGCAAGTGAGCGTGGAGCCATCGAGCTGCCCGAGAGAGCA 1417
Db 1201 GCTACGAGCTGCACCCGCAAGTGAGCGTGGAGCCATCGAGCTGCCCGAGAGAGCA 1260
QY 1418 GCTGGACCGTGAACGACATCCAGAGCTGGTGGCAAGCTGAAGTGGCGCCAGCCAGATCT 1477
Db 1261 GCTGGACCGTGAACGACATCCAGAGCTGGTGGCAAGCTGAAGTGGCGCCAGCCAGATCT 1320
QY 1478 ACCCGGCATCAAGTGGCGCAGCTGTGCAAGCTGTGCGCGCGCCCAAGGCCCTGACCG 1537
Db 1321 ACCCGGCATCAAGTGGCGCAGCTGTGCAAGCTGTGCGCGCGCCCAAGGCCCTGACCG 1380
QY 1538 ACATGTCGCCCTGACCGAGGAGCGAGCTGGAGCTGGCGCGAGAACCGGAGATCTGTC 1597
Db 1381 AGGTGATCCCCCTGACCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTGTA 1440
QY 1598 GCGAGCCCGTGCACGGCGTGTACTACGACCCCAAGAGGAGCTGGTGGCGGAGATCCAGA 1657
Db 1441 AGGAGCCCGTGCACGGCGTGTACTACGACCCCAAGAGGAGCTGGTGGCGGAGATCCAGA 1500
QY 1658 AGCAGGGCCAGCAGCTGACCTTACAGATCTACCAAGAGGAGCCCTTCAAGAACTGAAGA 1717
Db 1501 AGCAGGGCCAGCAGCTGACCTTACAGATCTACCAAGAGGAGCCCTTCAAGAACTGAAGA 1560
QY 1718 CCGGCAAGTACGCAAGATCGGACCCCGCCACACCAAGAGCTGAAGCAGCTGACCCAGG 1777
Db 1561 CCGGCAAGTACGCGGATCGCGGGCGGCCACCAACAGAGCTGAAGCAGCTGACCCAGG 1620

QY 1778 CCCTGCAGAAAGATCCCATGAGAGCATCGTATCTGGGGCAAGACCCCAAGTTCCGCC 1837
Db 1621 CCCTGCAGAAAGATCGACACCGAGAGCATCGTATCTGGGGCAAGATCCCAAGTTCAAGC 1680
QY 1838 TGCCCATCCAGAAAGAGACCTGGTGGAGACCTGGTGGACCGACTACTGGCAGGCGCACTGGA 1897
Db 1681 TGCCCATCCAGAAAGAGACCTGGTGGAGGCTGGTGGATGGATGCTGGCAGGCGCACTGGA 1740
QY 1898 TCCCGAGTGGAGTTCGTGAACACCCCGCTGGTGAAGCTGGTACCAAGTGGGAGA 1957
Db 1741 TCCCGAGTGGAGTTCGTGAACACCCCGCTGGTGAAGCTGGTACCAAGTGGGAGA 1800
QY 1958 AGGAGCCCATCATCGCGCCGAGACCTTCTACGTGGAGCGGCGCCCAACCGCGAGACA 2017
Db 1801 AGGAGCCCATCATCGCGCCGAGACCTTCTACGTGGAGCGGCGCCCAACCGCGAGACA 1860
QY 2018 AGATCGGCAAGCGCGCTTACGTGACCGGCGGCGGCGAGATCTGTGAGCCCTGACCG 2077
Db 1861 AGTGGGCAAGCGCGCTTACGTGACCGGCGGCGGCGGCGAGAGGTGGTGGATCGCGC 1920
QY 2078 AGACCAACCAAGAGACCGAGCTGAGGCCATCCAGCTGGCCCTGAGGACAGCGCA 2137
Db 1921 ACACCAACCAAGAGACCGAGCTGAGGCCATCCAGCTGGCCCTGAGGACAGCGCA 1980
QY 2138 GCGAGGTGAACATCGTGACCGCACAGCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCG 2197
Db 1981 TGGAGGTGAACATCGTGACCGCACAGCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCG 2040
QY 2198 ACAAGAGCGAGGAGCTGTGTGAACCAAGATCATCGACAGCTGATCAAGAGAGAGAGG 2257
Db 2041 ACAAGAGCGAGGAGCTGTGTGAACCAAGATCATCGACAGCTGATCAAGAGAGAGAGG 2100
QY 2258 TGTACCTGAGCTGGTGGCGCCCGCCCAAGAGGAGTGGGGCAACGAGCAGATPCGACAGC 2317
Db 2101 TGTACCTGAGCTGGTGGCGCCCGCCCAAGAGGAGTGGGGCAACGAGCAGTGGACAAGC 2160
QY 2318 TGGTGAAGCAAGGATCGCGCAAGTGTCTTCCCTGGACGGATCGATGCGGCATCTGTA 2377
Db 2161 TGGTGAAGCGCGCATCCGCAAGGTGTCTTCCCTGAACGGATCGATGCGGCATCTGTA 2220
QY 2378 TCTACCACTACATGAGCAGCTGTACGTGGCGAGCGCGGCGCTAGGATCGATTAAGAGC 2437
Db 2221 TCTACCACTACATGAGCAGCTGTACGTGGCGAGCGCGGCGGCGCTAGGATCGATTAAGAGC 2280
QY 2438 TTCCCGGGGCTAGCACCGGTGAATTC 2463
Db 2281 TTCCCGGGGCTAGCACCGGTGAATTC 2306

RESULT 5

ABK91616

ID - ABK91616 standard; DNA; 9166 BP.

XX ABK91616;

AC AC

XX XX

DT 14-AUG-2002 (first entry)

XX XX

DE Modified HIV protein-encoding plasmid DNA #168.

XX XX

KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX Synthetic.

OS OS

XX WO200232943-A2.

XX XX

XX 25-APR-2002.

XX XX

XX 14-AUG-2001; 2001WO-US25721.

XX XX

XX 14-AUG-2000; 2000US-225097P.

QY	818	AGATCGGCCCCCGAGAACCCCTACAAACACCCCGCTTTCGCATCAAGAAAGAGCAGCA	877
DB	661	AGATCGGCCCCCGAGAACCCCTACAAACACCCCGTTTCGCATCAAGAAAGAGCAGCA	720
QY	878	CCAAGTGGGGCAGAGCTGTGGACTTCGCGAGCTGAACAAGCGACCCAGGACTTCTGGG	937
DB	721	CCAAGTGGGGCAGAGCTGTGGACTTCGCGAGCTGAACAAGCGACCCAGGACTTCTGGG	780
QY	938	AGGTGCAGCTGGGCATCCGCCACCCCGCGCCTGAAGAAGAAAGAGCGTGACCGTGC	997
DB	781	AGGTGCAGCTGGGCATCCGCCACCCCGCGCCTGAAGAAGAAAGAGCGTGACCGTGC	840
QY	998	TGAGCTGGGGCAGCGCTTACTTTACAGGTGCCCCCTGGAGAGGACTTCGCGAAGTACACCG	1057
DB	841	TGAGCTGGGGCAGCGCTTACTTTACAGGTGCCCCCTGGAGAGGACTTCGCGAAGTACACCG	900
QY	1058	CCTTACCATCCCGACGATCAACAACGAGACCCCGGCATCCGCTACCACTACAACGTGC	1117
DB	901	CCTTACCATCCCGACGATCAACAACGAGACCCCGGCATCCGCTACCACTACAACGTGC	960
QY	1118	TGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGACTCTGG	1177
DB	961	TGCCCCAGGGCTGGAAGGGCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGACTCTGG	1020
QY	1178	AGCCCTTCGCGCCCGCAACCCCGAGATCGTGATCTACCA-----GGCCCCCTGTACG	1231
DB	1021	AGCCCTTCGCGCAAGCAGAACCCCGACATCGTGATCTACCACTACATGGAGCAGCTGTACG	1080
QY	1232	TGGGCAGCGACCTTGAGATCGGCCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCACC	1291
DB	1081	TGGGCAGCGACCTTGAGATCGGCCAGCACCGCAAGATCGAGGAGCTGCGCCAGCACC	1140
QY	1292	TGCTGGCTGGGGCTTCACACCCCGCACAGAAGCAGCACCAGAGAGCGCCCTTCTGT	1351
DB	1141	TGCTGGCTGGGGCTTCACACCCCGCACAGAAGCAGCACCAGAGAGCGCCCTTCTGT	1200
QY	1352	GGATGGCTACGAGCTGCAACCCCGCAAGTGGACCGTGCAGCCCATCGAGTGCCTCGAGA	1411
DB	1201	GGATGGCTACGAGCTGCACCCCGACAGTGGACCGTGCAGCCCATCATGCTGCCGAGA	1260
QY	1412	AGGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAATGGGCAGCC	1471
DB	1261	AGGAGCTGGACCGTGAACGACATCCAGAAGCTGGTGGGCAAGCTGAATGGGCAGCC	1320
QY	1472	AGATCTACCCCGGCATCAAGTGGCGCAGCTGTGCAAGCTGCTCGCGGGCGCAAGSCCC	1531
DB	1321	AGATCTACCCCGGCATCAAGTGAAGCAGCTGTGCAAGCTGCTCGCGGGCACCAGGCC	1380
QY	1532	TGACCGACATCTGTGCCCTTGACCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGGAGA	1591
DB	1381	TGACCGAGGTATCCCCCTGACCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGGAGA	1440
QY	1592	TCCTGCGCAGCCCGTGCAAGCGGTGTACTAGGACCCAGCAAGACCTGGTGGCCGAGA	1651
DB	1441	TCCTGAAGAGCCCGTGCAAGAGGTGTACTAGGACCCAGCAAGACCTGGTGGCCGAGA	1500
QY	1652	TCCAGAAGCAGGGCCACGACCACTACAGATCTACCAAGTCTACCAAGAGCCCTTCAAGAACC	1711
DB	1501	TCCAGAAGCAGGGCCAGGCGAGTGGACCTACCAAGTCTACCAAGAGCCCTTCAAGAACC	1560
QY	1712	TGAAGACCGCAAGTACGCCAAGATGCGCACCGCCCCACCAACAGCAGTGAAGCAGCTGA	1771
DB	1561	TGAAGACCGCAAGTACGCCCGCATCGCGGGCGGCCACCAACAGCAGTGAAGCAGCTGA	1620
QY	1772	CCGAGGCGGTGCAAGAGATCGCCATCGAGAGATCGTGATCTGGGGCAAGACCCCAAGT	1831
DB	1621	CCGAGGCGGTGCAAGAGGTAGCACCGAGAGATCGTGATCTGGGGCAAGATCCCAAGT	1680
QY	1832	TCCGCTTGCCCATCCAGAAGGAGACTGGGAGACCTGGTGACCGACTACTTGGCAGGCCA	1891
DB	1681	TCAAGCTGCCCATCCAGAAGGAGACTGGGAGGCGCTGGTGGATGGAGTACTTGGCAGGCCA	1740

Qy	1892	CCTGGATCCCGGAGTGGGAGTTCGTGTAACACACCCCCCCCTGGTGAAGCTGTGGTACCAGC	1951
Db	1741		
		CCTGGATCCCGGAGTGGGAGTTCGTGTAACACACCCCCCTGGTGAAGCTGTGGTACCAGC	1800
Qy	1952	TGGAGAGGAGGCCCATCATCTGGCGCCGAGACCTTCTACGTGGACGGGCCGCCCAACCGCG	2011
Db	1801		
		TGGAGAGGAGGCCCATCTGGTGGCGCCGAGACCTTCTACGTGGACGGCGCGCCCAACCGCG	1860
Qy	2012	AGACCAAGATCGCCACAGGCGGCTACGTGACCGACCGGGCGCGGCAGAAAGATCGTGAACC	2071
Db	1861		
		AGACCAAGCTGGGCAAGCGCGGTAGCTGACCGACCGGGCGCGCAGAAAGTGTGAGCA	1920
Qy	2072	TGACCGAGACACCAACCAAGAGACCGAGCTGCAGGCCATCCAGCTGBCCTCGCAGGACA	2131
Db	1921		
		TCGCCGACACCAACCAAGAGACCGAGCTGCAGGCCATCCACCTGGCCCTGCAGGACA	1980
Qy	2132	CGCGCAGGAGGTGAACATCTGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCC	2191
Db	1981		
		CGGCCCTGGAGGTGAACATCTGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCC	2040
Qy	2192	AGCCCGACAAGACGAGCGAGCGAGCTGGTGAACACAGATCATCGACGAGCTGATCAAGAAG	2251
Db	2041		
		AGCCCGACAAGAGGAGCGAGCTGGTGAAGCCAGATCATCGACGAGCTGATCAAGAAG	2100
Qy	2252	AGAAAGTGTACCTGAGCTGGTGGTGCCCGCCCAAGAGGCATCGGGGCAACGAGCAGATCG	2311
Db	2101		
		AGAAAGTGTACCTGGCTGGGTGCCCGCCCAAGAGGCATCGGGGCAACGAGCAGGTGG	2160
Qy	2312	ACAAGCTGTGAGCAAGGGCATCCGCAAGGTGCTGTTCTTGGACGGGCATCGATGGCGGCA	2371
Db	2161		
		ACAAGCTGGTGAGCGCGGCATCCGCAAGGTGCTGTTCTTGAACGGGCATCGATGGCGGCA	2220
Qy	2372	TCGTGATCTACCACTACATGGACGACCTGTAGCTGGGACGCGGGCCCTAGGATCGATT	2431
Db	2221		
		TCGTGATCTACCACTACATGGACGACCTGTAGCTGGGACGCGGGCCCTAGGATCGATT	2280
Qy	2432	AAAAGCTTCCCGGGCTAGCACCGGTGAATTC	2463
Db	2281	AAAAGCTTCCCGGGCTAGCACCGGTGAATTC	2312

DECEMBER 7

RESULT /
AAA70480
ID AAA70480 standard; DNA: 2300 BP.

AA
AC
AAA70480:XX
DT 28-NOV-2000 (first entry)

DE HTV ES(-) ProtMod RTopt YMWM coding sequence. XX

XX
KW HTV-1: AIDS: Gag-reverse transcriptase: vaccine: expression cassette: ss.

XX	OS	OS
	Human immunodeficiency virus type 1.	Synthetic.

XX
US
SYNCHETIC.

PN WO200039302-A2.

06-JUL-2000.

AA
PF 30-DEC-1999; 99WO-US31245.

AX 31-DEC-1998; 98US-0114495.

PR 01-DEC-1999; 99US-0168471.

PA (CHIR) CHIRON CORP.

Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H; PI
PI Greer C, Selby M, Walker C;

DR WPI; 2000-452400/39.

PT Expression cassettes encoding the human immunodeficiency virus (HIV)

PT Gag-containing polypeptide useful for vaccinating against HIV
PT infections and acquired immunodeficiency syndrome (AIDS) -
XX Claim 7; Fig 74; 391pp; English.

CC The present sequence is the coding sequence of a synthetic HIV
CC gag-reverse transcriptase expression cassette, FS(-).ProtMod.Rtopt_MMM.
CC The gag protein of HIV is needed for the assembly of virus-like
CC particles. In addition, the gag protein is involved in many stages of the
CC HIV life cycle, including assembly, virion maturation after particle
CC release and early post-entry steps in viral replication. The expression
CC cassette may be used for the recombinant expression of HIV
CC gag-polypeptides which may then be used to vaccinate against HIV
XX infection and acquired immunodeficiency syndrome (AIDS).

XX Sequence 2300 BP; 528 A; 754 C; 725 G; 293 T; 0 other;

Query Match 82.0%; Score 2019.2; DB 21; Length 2300;
Best Local Similarity 93.2%; Pred. No. 5.4e-245;
Matches 2150; Conservative 0; Mismatches 138; Indels 18; Gaps 3;

Qy 170 GGGCAAGGAGGGCCACACAGATGAAGGACTGCACCGAGCGCCAGCCAACTTCTTCGCG 229
Db 1 GGGCGCGAGGACACCAATGAAGATTGCATGAGAGACAGCTTAATTTCTTCGCG 60
Qy 230 AGGACTGGCTTCCCGCAGGCAAGGCCGCGAGTTCCCGAGGAGCAACCCGCGCA 289
Db 61 AGGACTGGCTTCTCGAGGCAAGGCCGCGAGTTTCAGCAGGAGGAGACCCGCGCA 120
Qy 290 ACAGCCCCACAGCCGCGAGCTGCAGTGGCGGGG-----ACAACCCCGCAGCGAGG 343
Db 121 ACAGCCCCACCGCGCGAGTGCAGTGTGGGGGCGGAGAACACAGCCTGAGCGAGG 180
Qy 344 CGGCGCGAGCGCAGGCGACCCCTG-----AACTTCCCGAGATCACCTGTGGCAGC 397
Db 181 CGGCGCGAGCGCAGGCGACCTGAGCTCACTTCCCGAGATCACCTGTGGCAGC 240
Qy 398 GCCCCTGGTGAGCATCAAGTGGCGGCCAGATCAAGGAGGCCCTGTGGACACCGCG 457
Db 241 GCCCCTGGTGACCATCAGATCGCGGCCAGCTCAAGGAGCGCTGTGACACCGCG 300
Qy 458 CCGAGCACCGCTGTGAGGAGATGACCTTGGCCCGGCAAGTGGAGCCCAAGATGATCG 517
Db 301 CCGAGCACCGCTGTGAGGAGATGAACCTTCCCGGCAAGTGGAGCCCAAGATGATCG 360
Qy 518 CGGCGATCGCGGCTTCAAGTGGCGCCAGTACGACAGATCCTGATCAGATCTGCG 577
Db 361 CGGCGATCGGGGCTTCAAGTGGCGGAGTAGGACACAGATCCCGTGGAGATCTGCG 420
Qy 578 GCAAGAGGCCATCGGCACCGTGTGATCGGCCCGCACCCCGTGAAACATCATCGGCCGCA 637
Db 421 GCCAAGGCCATCGGCACCGTGTGTTGGCGCCCGCACCCCGTGAAACATCATCGGCCGCA 480
Qy 638 ACATGCTGACCCAGCTGGGCTGCACCTGAACTTCCCATCAGCCCATCAGACCGTGC 697
Db 481 ACCTGCTGACCCAGATCGGCTGCACCTGAACTTCCCATCAGCCCATCAGACCGTGC 540
Qy 698 CCGTGAAGCTGAAGCCCGGCATGGACGGCCCAAGTGAAGCGTGGCCCTGACCGAGG 757
Db 541 CCGTGAAGCTGAAGCCGGGATGGAGGCCCAAGTGAAGCGTGGCCCTGACCGAGG 600
Qy 758 AGAGATCAAGCCCTGACCCGATCTCGAGGAGATGGAGAGGAGGGCAAGATCAACA 817
Db 601 AGAGATCAAGCCCTGTGTGAGATCTGCACCGAGATGGAGAGGAGGGCAAGATCAACA 660
Qy 818 AGATGGCCCGAGAACCCCTACACACCCCGTGTTCGCCATCAAGAGAGAGACACA 877
Db 661 AGATGGCCCGAGAACCCCTACACACCCCGTGTTCGCCATCAAGAGAGAGAGACACA 720
Qy 878 CCAAGTGGCGCAAGCTGTGTGACTTCCCGAGCTGAACAAGCGCACCGAGCTTCTGGG 937
Db 721 CCAAGTGGCGCAAGCTGTGTGACTTCCCGAGCTGAACAAGCGCACCGAGCTTCTGGG 780

Qy 938 AGGTGAGCTGGGCTATCCCGCCCGCGCTGAAGAAGAGAGAGCGTGACCGTGC 997
Db 781 AGGTGAGCTGGGCTATCCCGCCCGCGCTGAAGAAGAGAGAGCGTGACCGTGC 840
Qy 998 TGGAGCTGGGCGAGCGCTACTTACGCGTCCCGTGCAGGAGACTTCCGCAAGTACACCG 1057
Db 841 TGGAGCTGGGCGAGCGCTACTTACGCGTCCCGTGCAGGAGACTTCCGCAAGTACACCG 900
Qy 1058 CTTTACCATTCCCGAGCATCAACAGCAGACCCCGGCTACCGTACCGTACAGCTGC 1117
Db 901 CTTTACCATTCCCGAGCATCAACAGCAGACCCCGGCTACCGTACCGTACAGCTGC 960
Qy 1118 TGCCCCAGGCTGGAAGGCGAGCCCGAGCATCTTCAGAGCAGCATGACCAAGATCCTGG 1177
Db 961 TGCCCCAGGCTGGAAGGCGAGCCCGGCTACCTTCAGAGCAGCATGACCAAGATCCTGG 1020
Qy 1178 AGCCCTTCGCGCCCGCAACCCCGAGATCGTATCTACAGGCGCCCTGTAGTGGGCA 1237
Db 1021 AGCCCTTCGCGAGGAGAACCCCGCATCTATCTACAGGCGCCCTGTAGTGGGCA 1080
Qy 1238 GCGAGCTGGAGATCGGCGAGCAGCCGCGCAAGATCGAGGCTGCGCAAGCAGCTCTGC 1297
Db 1081 GCGAGCTGGAGATCGGCGAGCAGCCGCGCAAGATCGAGGCTGCGCGCAGCAGCTCTGC 1140
Qy 1298 GCTGGGGCTTACACACCCCGCAGACAGCAGCAGAGGAGCCCGCTTCTCTGTGATGG 1357
Db 1141 GCTGGGGCTTACACACCCCGCAGACAGCAGCAGAGGAGCCCGCTTCTCTGCCAT-- 1198
Qy 1358 GCTAGAGCTGCACCCCGCAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGA 1417
Db 1199 ----CGAGCTGCACCCCGCAAGTGGACCGTGCAGCCCATCTGTGCCCGAGAGGACA 1254
Qy 1418 GCTGAGCCTGAACACACATCCAGAGCTGGTGGCAAGCTGAACCTGGGCGCAGCCAGTCT 1477
Db 1255 GCTGAGCCTGAACACACATCCAGAGCTGGTGGCAAGCTGAACCTGGGCGCAGCCAGTCT 1314
Qy 1478 ACCCGGGCTCAAGTGGCGCCAGCTGTGCAAGCTGTGCGCGCGCGCAAGGCGCTGACCG 1537
Db 1315 AGCGCGGCTCAAGTGAAGTGAAGCAGCTGCAAGCTGCTGCGCGCAGCAGGCGCTGACCG 1374
Qy 1538 ACATCGTGGCCTGACCGAGGAGCGGAGCTGGAGTGGCCGAGAACCCGAGAGTCTCTGC 1597
Db 1375 AGGTGATCCCGTGCAGGAGGCGGAGCTGGAGTGGCGCGAGAACCCGAGAGTCTCTGA 1434
Qy 1598 GCGAGCCGCTGCACCGCGTGTACTAGACCCCGCAGAACACCTGTGGCCGAGATCCACA 1657
Db 1435 AGGAGCCCGTGCACCGAGGTGTACTAGACCCCGCAGAACACCTGTGGCGAGATCCACA 1494
Qy 1658 AGCAGGGCCACGACAGTGGAGCTTACAGATCTACAGAGCGCTTCAAGAACCTGAAGA 1717
Db 1495 AGCAGGGCCAGGCGCAGTGGACCTACAGATCTACAGAGCGCTTCAAGAACCTGAAGA 1554
Qy 1718 CCGGCAAGTACGCAAGATGCGCCACCGCCCAACACAGCAGTGAAGCAGTGAAGCGAGG 1777
Db 1555 CCGGCAAGTACGCGCGCATGCGCGCGCCCAACACAGCAGTGAAGCAGTGAAGCGAGG 1614
Qy 1778 CCGTCAAGAAATCGCCATGGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCGCGC 1837
Db 1615 CCGTCAAGAAATGGAGCAGCAGCAGAGCATCTGATCTGGGGCAAGATCCCAAGTCAAGC 1674
Qy 1838 TGCCCATCCAGAGGAGACCTGGGAGACCTGGTGGACCGAGCTACTTGGAGCGCCAGCTGA 1897
Db 1675 TGCCCATCCAGAGGAGACCTGGGAGGCTGTGTGATGGATGAGTACTGGCAGGCGACCTGA 1734
Qy 1898 TCCCGAGTGGAGTTCGTGAACACCCCGCTGGTGAAGCTGTGTACAGTGGGAGA 1957
Db 1735 TCCCGAGTGGAGTTCGTGAACACCCCGCTGGTGAAGCTGTGTACAGTGGGAGA 1794
Qy 1958 AGGAGCCCATCATCGCGCGCAGACCTTCTAGTGGAGCGCGCCCAACCCCGAGAGACA 2017
Db 1795 AGGAGCCCATCATCGTGGCGCGCAGACCTTCTAGTGGAGCGCGCCCAACCCCGAGAGACA 1854
Qy 2018 AGATCGGCAAGGCGCGCTACGTGACCGAGCGCGCGCAGAGATCGTGTAGCCTGACCG 2077

1855 AGCTGGCAAGCGCGCTACGTGACCGACCGCGCGCGCAGAGGTGGTGAGCATCGCG 1914
2078 AGACCAACCAAGACAGCGAGCTGCGAGCGCATCCAGCTGGCCCTGCGAGCAGCGGCA 2137
1915 ACACCAACCAAGACAGCGAGCTGCGAGCGCATCCAGCTGGCCCTGCGAGCAGCGGCG 1974
2138 GCGAGGTGAACATCGTGACCGACAGCGAGTACGCGCTGGCGCATCATCGAGCGCGAGCGCG 2197
1975 TGAGGTGAACATCGTGACCGACAGCGAGTACGCGCTGGCGCATCATCGAGCGCGAGCGCG 2034
2198 ACAAGAGCGAGCGAGCTGGTGAACACAGATCATCGAGCGAGCTGATCAAGAAGGAGAAG 2257
2035 ACAAGAGCGAGCGAGCTGGTGAACACAGATCATCGAGCGAGCTGATCAAGAAGGAGAAG 2094
2258 TGACCTGAGCTGGTGCCCGCCCAAGGCGATCGCGCGAGCGCGCGAGCGAGCGATCGACAAGC 2317
2095 TGTACCTGGCTGGTGCCCGCCCAAGGCGATCGCGCGAGCGCGAGCGAGCGAGCGATCGACAAGC 2154
2318 TGTGAGCAAGGCGATCGCGCAAGGTGCTGTTCTGAGCGCGATCGATGCGCGGCGATCGTGA 2377
2155 TGTGAGCGCGCGATCGCGCAAGGTGCTGTTCTGAGCGCGATCGATGCGCGGCGATCGTGA 2214
2378 TGTACCAAGTACATGAGCGAGCTGTACGTGGCGAGCGCGCGCGCGCGAGCGATCGATTAAGC 2437
2215 TGTACCAAGTACATGAGCGAGCTGTACGTGGCGAGCGCGCGCGCGCGAGCGATCGATTAAGC 2274
2438 TTTCCCGGGCTAGCACCGGTGAATTC 2463
2275 TTTCCCGGGCTAGCACCGGTGAATTC 2300

RESULT 8

ABK91622

ID ABK91622 standard; DNA; 9788 BP.

XX

AC ABK91622;

XX

14-AUG-2002 (first entry)

XX

DE Modified HIV protein-encoding plasmid DNA #174.

XX

HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX

OS Synthetic.

XX

XX WO20232943-A2.

PN

XX

XX 25-APR-2002.

XX

14-AUG-2001; 2001WO-US25721.

XX

14-AUG-2000; 2000US-225097P.

PR

14-NOV-2000; 2000US-252115P.

PR

28-MAR-2001; 2001US-279257P.

XX

(USSH) US DEPT HEALTH & HUMAN SERVICES.

PA

(CHAD/) CHADABARTI B K.

XX

Nabel GJ, Huang Y;

XX

WPI; 2002-452382/48.

XX

New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
its encoded protein, useful as vaccines for genetic or protein
immunisation for acquired immunodeficiency syndrome or human
immunodeficiency virus infection

XX

Disclosure; Page 785-788; 794pp; English.

PS

The invention relates to a nucleic acid molecule encoding a modified HIV

XX

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

XX

SQ Sequence 9788 BP; 2377 A; 2817 C; 2695 G; 1899 T; 0 other;

Query Match

Best Local Similarity 82.0%; Score 2019.2; DB 24; Length 9788;

Matches 2190; Conservative 0; Mismatches 178; Indels 24; Gaps 4;

QY

14 TGCCCGAGGCGCATGAGCGAGCGCCACAGCGCCAACTCTGATGCGAGCGCAGCAACTTCA 73
Db TGCCCGAGGCGCATGAGCGAGCGCCAACTCTGATGCGAGCGCAGCAACTTCA 3020

QY

74 AGGCGCCCAAGCGCATCATCAAGTGTCTCACTCGCGCAAGGAGGCGCACATCGCCGCA 133
Db AGGCGCCCAAGCGCATCATCAAGTGTCTCACTCGCGCAAGGAGGCGCACATCGCCGCA 3080

QY

134 ACTGCCCGCGCCCGCCCAAGAGGGCTGCTGGAAGTGCAGGAGGAGGCGCCACAGATGA 193
Db ACTGCCCGCGCCCGCCCAAGAGGGCTGCTGGAAGTGCAGGAGGAGGCGCCACAGATGA 3140

QY

194 AGGACTCACCGAGCGCGCCAGGCGCAACTTCTTCCCGAGGAGCTTGGCTTCCCCAGGGCA 253
Db AGGACTCACCGAGCGCGCGCCAGGCGCAACTTCTTCCCGAGGAGCTTGGCTTCCCCAGGGCA 3194

QY

254 AGGCGCGCGAGTTCGCCAGCGAGCGAGAACCGCGCCACACGCCCCACAGCGCGCGAGCTGC 313
Db AGGCGCGCGAGTTCGCCAGCGAGCGAGAACCGCGCCACACGCCCCACAGCGCGCGAGCTTC 3254

QY

314 AGGTGCGCGG-----CGACAACCCCGCGAGCGAGCGCGCGCGCGCGCGCGCGCA--- 364
Db AGGTGCGCGG-----CGACAACCCCGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCA--- 3314

QY

365 ---CCCTGAATTCCTCCCGAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAGTGG 421
Db TATCTTTAGCTTCTCCCTCAGATCACTCTTTGGCAGCACCCCTCTCAACAATAAGATAG 3374

QY

422 GCGCGCAGATCAAGGAGCGCGCTGCTGGACACCGCGCGCGAGCACCGCTGCTGGAGGAGA 481
Db GCGCGCAGATCAAGGAGCGCGCTGCTGGACACCGCGCGCGAGCACCGCTGCTGGAGGAGA 3434

QY

482 TGAGCCTTGCCTGCGCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGGGGCTTCATCAAG 541
Db TGAGCCTTGCCTGCGCGCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGGGGCTTCATCAAG 3494

QY

542 TGCGCCAGTACGACAGATCCTGATCGAGATCTCGCGCAAGAGGCCATCGGCACCGTGC 601
Db TGCGCCAGTACGACAGATCCTGATCGAGATCTCGCGCAAGAGGCCATCGGCACCGTGC 3554

QY

602 TGATCGCGCGCGCGCGCGTGAACATCATCGCGCGCACATCTGACCCAGCTGGGCTGCA 661
Db TGATCGCGCGCGCGCGCGTGAACATCATCGCGCGCACATCTGACCCAGCTGGGCTGCA 3614

QY

662 CCCTGAATTCCTCCCATCAGCGCCCATCGAGACCGTGCCTGCGCGTGAAGCTGAAGCCCGCATGG 721
Db CCCTGAATTCCTCCCATCAGCGCCCATCGAGACCGTGCCTGCGCGTGAAGCTGAAGCCCGCATGG 3674

QY

722 ACGCCCCCAAGGTGAAGCAGTGGCCCTTGACCGAGAGAGATCAAGGCCCTGACCGCCA 781
Db ACGCCCCCAAGGTGAAGCAGTGGCCCTTGACCGAGAGAGATCAAGGCCCTGACCGCCA 3734

QY

782 TCTCGAGGAGATGGAGAGGAGGAGATCACCAGATCGCGCGCGAGAACCCCTACA 841
Db TCTCGAGGAGATGGAGAGGAGGAGATCACCAGATCGCGCGCGAGAACCCCTACA 3794

QY

QY 842 ACACCCCGTGTTCGCCATCAAGAAAGAGACAGACACCAAGTGGCGCAAGCTGGTGGACT 901
Db ACACCCCGTGTTCGCCATCAAGAAAGAGACAGACACCAAGTGGCGCAAGCTGGTGGACT 3854
QY 902 TCCGCGAGCTCAAGACGCGACCCAGGACTTCTGGAGGTCAGCTGGGCATCCCCACCC 961
Db TCCGCGAGCTCAAGACGCGACCCAGGACTTCTGGAGGTCAGCTGGGCATCCCCACCC 3914
QY 962 CCGCGGCTGAAGAAAGAGAGCGGTGACCGTGTGGAGCTGGGCGACGCCCTACTTCA 1021
Db CCGCGGCTGAAGACGAGAAAGAGCGGTGACCGTGTGGAGCTGGGCGACGCCCTACTTCA 3974
QY 1022 CGTGGCCCTGGAGAGACTTCCGCAAGTACACCGCCCTTCAACATCCCCAGCATCAACA 1081
Db CGTGGCCCTGGAGAGACTTCCGCAAGTACACCGCCCTTCAACATCCCCAGCATCAACA 4034
QY 1082 ACAGACCCCGGATCCGCTACCAAGTACACGTCGTCGCCAGGCGTGGAGGCGACGCC 1141
Db ACAGACCCCGGATCCGCTACCAAGTACACGTCGTCGCCAGGCGTGGAGGCGACGCC 4094
QY 1142 CCAGCATCTTCCAGAGCAGCATGACCAAGATCCCTGGAGCCCTTCCGCGCCCGCAACCCCG 1201
Db CCAGCATCTTCCAGTGCAGCATGACCAAGATCCCTGGAGCCCTTCCGCGCGCAAGACCCCG 4154
QY 1202 AGATCGTGTACTACA-----GGCCCCCTGTAGTGGGCGAGCACCTGGAGATCGGCC 1255
Db AGATCGTGTACTACAGTATACATGACCAACCTGTACCTGGGCGAGCACCTGGAGATCGGCC 4214
QY 1256 AGACCGCGCAAGATCGAGGAGCTGGCGACGACCTGCTGCGCTGGGGCTTCAACACCC 1315
Db AGACCGCGCAAGATCGAGGAGCTGGCGACGACCTGCTGCGCTGGGGCTTCAACACCC 4274
QY 1316 CCGAAGAAGACACCAAGAGAGCCCTTCTGTGTGGATGGGCTACGAGCTGCACCCCG 1375
Db CCGAAGAAGACACCAAGAGAGCCCTTCTGTGTGGATGGGCTACGAGCTGCACCCCG 4334
QY 1376 ACAAGTGCAGCTGACGCCATCGAGCTGCCGAGAGAGAGAGTGGACCGTGAACGACA 1435
Db ACAAGTGCAGCTGACGCCATCGAGCTGCCGAGAGAGAGAGTGGACCGTGAACGACA 4394
QY 1436 TCCAGAAGCTGTGGGCAAGCTGAACCTGGCGACCGCAGATCTACCCGCGATCAAGTGC 1495
Db TCCAGAAGCTGTGGGCAAGCTGAACCTGGCGACCGCAGATCTACCCGCGATCAAGTGC 4454
QY 1496 GCCAGCTGTGCAAGCTGCTGCGCGCGCCCAAGGCCCTGACGACATGTCGCCCTGACCG 1555
Db GCCAGCTGTGCAAGCTGCTGCGCGCGCCCAAGGCCCTGACGAGGTGTCGCCCTGACCG 4514
QY 1556 AGGAGCGGAGCTGGAGCTGGCGGAGACCGGAGATCTCGCGGAGCCGCTGCACGCG 1615
Db AGGAGCGGAGCTGGAGCTGGCGGAGACCGGAGATCTCGCGGAGCCGCTGCACGCG 4574
QY 1616 TGTACTACGACCCCGACGAGCTGTGGCGGAGATCCAGAAAGCAGGCGCCACACCACT 1675
Db TGTACTACGACCCCGACGAGCTGTGGCGGAGATCCAGAAAGCAGGCGCCACACCACT 4634
QY 1676 GGACCTACCAAGATCTACGAGAGCCCTTCAAGAACCTGAAAGACCGGCAAGTACGCCAAGA 1735
Db GGACCTACCAAGATCTACGAGAGCCCTTCAAGAACCTGAAAGACCGGCAAGTACGCCA 4694
QY 1736 TGGCCACCGCCACACAGAGCTGAAGAGCTGAGCGGAGCGCTGCAGAAATCGCCA 1795
Db TGGCCACCGCCACACAGAGCTGAAGAGCTGAGCGGAGCGCTGCAGAAATCGCCA 4754
QY 1796 TGGAGAGCATCTGTATCTGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAAAGAGA 1855
Db TGGAGAGCATCTGTATCTGGGCAAGACCCCAAGTTCCAGCTGCCATCCAGAAAGAGA 4814
QY 1856 CCTGGGAGACTGTGTGACCGACTTACCTGGCAGGCGACCTGGATCCCGAGTGGAGTTCG 1915
Db CCTGGGAGGCTGTGTGACCGAGTACTTGGCAGGCGACCTGGATCCCGAGTGGAGTTCG 4874
QY 1916 TGAACACCCCGCTGTGAGCTGTGGTACCAGCTGGAGAGAGGCCATCATCGCGC 1975

Db 4875 TGAACACCCCGCTGTGAGCTGTGGTACCAGCTGAGAAAGAGCCCATCATCGCGC 4934
QY 1976 CCGAGACCTTCTACGTGGACGCGCGCCACCAACCGGAGACCAAGATCGGAAGCGCGCT 2035
Db CCGAGACCTTCTACGTGGACGCGCGCGCCACCAACCGGAGACCAAGATCGGAAGCGCGCT 4994
QY 2036 ACGTGACCGACCGGCGCGGAGAGATCGTGAGCTGTACCGGAGACCAACCAAGAGAGA 2095
Db ACGTGACCGACCGGCGCGGAGAGATCGTGAGCTGTACCGGAGACCAACCAAGAGAGA 5054
QY 2096 CCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCGTGA 2155
Db CCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGTGGAGGTGAACATCGTGA 5114
QY 2156 CCGACAGCAGTACGCTTGGGCGATCCAGGCGCCCGCCAGAGAGCGAGAGCGAGC 2215
Db CCGACAGCAGTACGCTTGGGCGATCCAGGCGCCCGCCAGAGAGCGAGAGCGAGC 5174
QY 2216 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGGTGC 2275
Db TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGGTGC 5234
QY 2276 CCGCCCAAGAGGATCGCGGCAACGAGCAGATCGACAAGCTGCTGAGCAAGGCGCATCC 2335
Db CCGCCCAAGAGGATCGCGGCAACGAGCAGGTGGACGCGCTGCTGAGCGCGCGCATCC 5294
QY 2336 GCAAGGTGCTTCTCTGACCGCATCGATGGGCGATCGTGATCTACAGTA 2387
Db GCAAGGTGCTTCTCTGACCGCATCGATGGGCGCGATCGACAAGGCGCGAGAGCGAGAG 5346

RESULT 9

ABK91617

ID ABK91617 standard; DNA; 9169 BP.

XX ABK91617;

XX AC

XX 14-AUG-2002 (first entry)

XX DE Modified HIV protein-encoding plasmid DNA #169.

XX KW HIV; human immunodeficiency virus; gene; ds; circular; anti-HIV;
XX KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
XX KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX OS Synthetic.

XX PN WO200232943-A2.
XX XX

XX PD 25-APR-2002.

XX XX 14-AUG-2001; 2001WO-US25721.

XX XX 14-AUG-2000; 2000US-225097P.

XX PR 14-NOV-2000; 2000US-252115P.

XX PR 28-MAR-2001; 2001US-279257P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PA (CHAD/) CHADRABARTI B K.

XX PI Nabel GJ, Huang Y;

XX XX WPI; 2002-452382/48.

XX PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
XX PT its encoded protein, useful as vaccines for genetic or protein
XX PT immunisation for acquired immunodeficiency syndrome or human
XX PT immunodeficiency virus infection

XX PS Disclosure; Page 769-772; 794pp; English.
XX CC The invention relates to a nucleic acid molecule encoding a modified HIV

Db 4883 TGAACACCCCCCTGGTGAAGCTGTGGTACCAAGTGGAGAGGAGCCCATCATCGGG 4942
QY 1976 CCGAGACCTTCTAGCTGACGGCGCCGACACCGGAGACCAAGATCGGCAAGCGCGCT 2035
Db 4943 CCGAGACCTTCTAGCTGACGGCGCCGACACCGGAGACCAAGATCGGCAAGCGCGCT 5002
QY 2036 ACCTGACCGACCGGGCGCGGAGAGATCGTGAGCTGACCGAGACCAACCAACCAAGA 2095
Db 5003 ACCTGACCGACCGGGCGCGGAGAGATCGTGAGCTGACCGAGACCAACCAACCAAGA 5062
QY 2096 CCGAGCTGACGGCATCCAGCTGGCCCTGAGGACAGAGCGGAGAGGTGAACATCGTGA 2155
Db 5063 CCGAGCTGACGGCATCCACCTGGCCCTGAGGACAGAGCGGCTGAGAGTGAACATCGTGA 5122
QY 2156 CCGAGACGACGTAGCCCTGGGCTATCCAGAGCCGACGACCAAGAGCGAGCGAGC 2215
Db 5123 CCGAGACGACGTAGCCCTGGGCTATCCAGAGCCGACGACCAAGAGCGAGCGAGC 5182
QY 2216 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGTGTACCTGAGCTGGGTGC 2275
Db 5183 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGTGTACCTGAGCTGGGTGC 5242
QY 2276 CCGCCCAACAGGCGATCGGCGGACAGAGAGATCGACAAAGCTGGTGAAGGCGCATCC 2335
Db 5243 CCGCCCAACAGGCGATCGGCGGACAGAGAGTGGAGCGCTGGTGAAGGCGCATCC 5302
QY 2336 GCAAGGTGCTTCTTGGAGCGGATCGATGCGGCGATCGTATACAGTA 2387
Db 5303 GCAAGGTGCTTCTTGGAGCGGATCGATGCGGCGATCGTATACAGTA 5354

RESULT 10
ABK91612
ID ABK91612 standard; DNA; 9194 BP.
XX AC ABK91612;
XX 14-AUG-2002 (first entry)
DE Modified HIV protein-encoding plasmid DNA #164.
XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
OS Synthetic.
XX WO200232943-A2.
XX 25-APR-2002.
XX 14-AUG-2001; 2001WO-US25721.
XX 14-AUG-2000; 2000US-225097P.
PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADABARTI B K.
XX Nabel GJ, Huang Y;
PI WPI; 2002-452382/48.
XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX Disclosure; Page 753-756; 794pp; English.
XX The invention relates to a nucleic acid molecule encoding a modified HIV

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX

SQ Sequence 9194 BP; 2232 A; 2672 C; 2523 G; 1767 T; 0 other;
Query Match 81.3%; Score 2001.8; DB 24; Length 9194;
Best Local Similarity 91.3%; Pred. No. 6.8e-243;
Matches 2186; Conservative 0; Mismatches 187; Indels 22; Gaps 5;

QY 14 TGGCGAGCGCATGAGCCAGG---CCAGCAGCGCCCAACATCCTGATGAGCGCAGCAACT 70
Db 2967 TGGCGAGCGCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGAGCGCGCAACT 3026
QY 71 TCAGGGCCCCAAGCGCATCATCAAGTCTTCAACTGGCGCAAGAGGGCCACATCGGCC 130
Db 3027 TCCGCAACACAGCGCAAGATCGTGAAGTCTTCAACTGCGGCAAGAGGGCCACACCGCC 3086
QY 131 GCACTGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACACCA 190
Db 3087 GCACTGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACACCA 3146
QY 191 TGAAGGACTGACCGAGCGCCAGGCAACTTCTTCCGCGAGGACCTGGCTTCCCGCAGG 250
Db 3147 TGAAGGACTGACCGAGCGGACAGGCTAA-TTTTATGGAAGATCTGGCTTCCACAG 3205
QY 251 GCAAGGCCCGGAGTTCCCGAGCGAGCAGAACCGCGCCCAACAGCCCCACCGCCGAGC 310
Db 3206 GGAAGGCCAGGGAATTTTCTTACAGAGCAGACAGACCAACAGCCCCACCAAGAGAGC 3265
QY 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGCGCGCGCCGAGCGCCAGGCA 364
Db 3266 TGCAGGTGCGCGGAGAGACAACTCCCTCTCAGAACAGGAGCGCATGAGCAAGGAA 3325
QY 365 -----CCCTGAACCTTCCCCAGATCACCTTGTGACGCGCCCTGGTGTGACATCAAGG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTCTTGGCAGCGACCCCTGCTCACAATAAGA 3385
QY 419 TGGCGGCGCAGATCAAGAGGCGCTGCTGGACACCGCGCGCGAGACACCGTGTGGAGG 478
Db 3386 TAGGGGCGCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGGAGACACCGTGTGGAGG 3445
QY 479 AGATGAGCTGCGCGCAAGTGAAGCCCAAGATGATCGGCGGATCGGCGGCTTCATCA 538
Db 3446 AGATGAGCTGCGCGCGCGCTGGAAGCCCAAGATGATCGGCGGATCGGCGGCTTCATCA 3505
QY 539 AGGTGCGCCAGTACGACAGATCTCTGATCGAGATCTGCGGCAAGAGCCATCGGCACCG 598
Db 3506 AGGTGCGCCAGTACGACAGATCTCTGATCGAGATCTGCGGCGCACAAGCCATCGCACCG 3565
QY 599 TGCTGATCGGCGCCACCGCGTGAACATCATCGGCGCGCAACATGCTGACCGAGTGGCT 658
Db 3566 TGCTGATCGGCGCCACCGCGTGAACATCATCGGCGCGCAACATGCTGACCGAGTGGCT 3625
QY 659 GCACCTGAACTTCCCATGAGCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCGCA 718
Db 3626 GCACCTGAACTTCCCATGAGCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCGCA 3685
QY 719 TGGACGCGCCCAAGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 3686 TGGACGCGCCCAAGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGTTGG 3745
QY 779 CCATCTCGGAGGAGATGGAGAGGAGGCAAGATCAACCAAGATCGGCGCGGAGAACCCCT 838
Db 3746 AGATCTCGGAGGAGATGGAGAGGAGGCAAGATCAACCAAGATCGGCGCGGAGAACCCCT 3805

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
 CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
 CC genetic or protein immunisation to a host, respectively. In particular
 CC these are useful for ameliorating the symptoms of acquired
 CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
 CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
 CC compositions are useful for treating or preventing HIV infections or
 CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
 CC plasmid DNA sequences of the invention.

XX Sequence 9194 BP: 2232 A; 2671 C; 2524 G; 1767 T; 0 other;

Seq Match 81.2%; Score 2000.2; DB 24; Length 9194;
 Best Local Similarity 91.2%; Pred. No. 1.1e-242;
 Matches 2185; Conservative 0; Mismatches 188; Indels 22; Gaps 5;

QY 14 TGGCCGAGGCCATGAGCCAGG---CCACAGCGGCCACATCCTGATGAGCGCAGCAACT 70
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 71 TCAAGGGCCCAAGCGCATCATCAAGTGCCTTCAACTGCGCAAGGAGGGCCACATCGCCC 130
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3027 TCGCCAACACGCGCAAGATCGTGAAGTGCCTTCAACTGCGCAAGGAGGGCCACACCGCCC 3086
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 131 GCAACTGCCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA 190
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3087 GCAACTGCCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA 3146
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 191 TGAAGGACTGACGAGCGCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA 250
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3147 TGAAGGACTGACGAGCGCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA 3205
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 251 GCAAGGCGCGGAGTTCGCCAGCGAGCAGACCGCGCCACAGCGCCACACCGCGCAGC 310
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3206 GGAAGGCGAGGATTTCTTCAGAGCAGACCGAGCCACAGCGCCACACCGCGCAGC 3265
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 311 TGCAGGTGCGGG-----CGACAAACCCCGCAGCGAGGCGCGCGAGCGCGCGCAGGCA 364
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3266 TGCAGGTGCGGG-----CGACAAACCCCGCAGCGAGGCGCGCGAGCGCGCGCAGGCA 364
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 365 -----CCCTGAATTCCTCCCGAGCAGTCCCTGTCAGAGGAGGAGGCGCGCGCAGGCA 3325
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3326 CTGTATCTTTAGCTTCCCTTCAGATCACTCTTTGGCAGCAGCCCTGCTCACAATAAGA 3385
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 419 TGGCGGCGCAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGAGCAGCAGCGTGTGAGG 478
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3386 TAGGGGCGCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGAGCAGCAGCGTGTGAGG 3445
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 479 AGATGAGCTTGCCTGGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGGCTTCATCA 538
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3446 AGATGAACCTTGCCTGGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGGCTTCATCA 3505
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 539 AGTGGCGCAGTACGACACATCTGATCGAGATCTGCGCGCAAGAGGCCATCGGACCG 598
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3506 AGTGGCGCAGTACGACACATCTGATCGAGATCTGCGCGCAAGAGGCCATCGGACCG 3565
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 599 TGCTGATGCGGCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCGAGTGGGCT 658
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3566 TGCTGATGCGGCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCGAGTGGGCT 3625
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 659 GCACCTGGAATCTCCCATCAGCCCATCGAGACCGTGGCCCGTGAAGCTGAAGCCCGGCA 718
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3626 GCACCTGGAATCTCCCATCAGCCCATCGAGACCGTGGCCCGTGAAGCTGAAGCCCGGCA 3685
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 719 TGCACGCGCCCAAGGTCAAGCACTGGCCCTGACCGAGGAGAGATCAAGCCCTGACCG 778
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3686 TGCACGCGCCCAAGGTCAAGCACTGGCCCTGACCGAGGAGAGATCAAGCCCTGACCG 3745
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 779 CCATCTCGGAGGAGTGGAGAGGAGGCGCAAGATCAACAGATCGGCGCGCGAGAACCCCT 838
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
 QY 3746 AGATCTGCACCGAGATGGAGAGGAGGCGCAAGATCAACAGATCGGCGCGCGAGAACCCCT 3805
 Db TGGCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026

QY 839 ACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
 Db ACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGG 3865
 QY 899 ACTTCGCGAGCTGAACAAAGCGCACCCAGAGCTTCTTGGAGAGTGCAGCTGGGCACTCCGCC 958
 Db ACTTCGCGAGCTGAACAAAGCGCACCCAGAGCTTCTTGGAGAGTGCAGCTGGGCACTCCGCC 3925
 QY 959 ACCCGCGCGGCTGAAGAAGAAGAGACGCTGACCGTCTGGACGTGGGCGAGCGCTACT 1018
 Db ACCCGCGCGGCTGAAGAAGAAGAGACGCTGACCGTCTGGACGTGGGCGAGCGCTACT 3985
 QY 1019 TCAGCGTCCCGCTGGACGAGGACTTCCGCAAGTACACACGCTTCCACCATCCCCAGCATCA 1078
 Db TCAGCGTCCCGCTGGACGAGGACTTCCGCAAGTACACACGCTTCCACCATCCCCAGCATCA 4045
 QY 1079 ACAACGAGACCCCGCGCATCCCTACCAAGTACAAAGTCTGCTCCCGCAGGCGCTGGAAGGCA 1138
 Db ACAACGAGACCCCGCGCATCCCTACCAAGTACAAAGTCTGCTCCCGCAGGCGCTGGAAGGCA 4105
 QY 1139 GCGCCGAGCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCGCAACC 1198
 Db GCGCCGAGCTTCCAGAGTGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCGCAACC 4165
 QY 1199 CCGAGATCGTATCTACCA-----GGCCCCCTCTACGTGGGCGAGCGCTGGAGATCG 1252
 Db CCGAGATCGTATCTACCA-----GGCCCCCTCTACGTGGGCGAGCGCTGGAGATCG 4225
 QY 1253 GCCAGCAGCGCGCGCATGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGCGCTTCACCA 1312
 Db GCCAGCAGCGCGCATGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGCGCTTCACCA 4285
 QY 1313 CCGCAAGTGGACCGTGCAGCCCATCGAGTGCAGGAGGAGCGCGCGAGGAGCGCTGGAACG 1432
 Db CCGCAAGTGGACCGTGCAGCCCATCGAGTGCAGGAGGAGCGCGCGAGGAGCGCTGGAACG 4405
 QY 1433 ACATCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGAGCGAGATCTACCCCGCATCAAGG 1492
 Db ACATCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGAGCGAGATCTACCCCGCATCAAGG 4465
 QY 1493 TGGCGCAGCTGTCAAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1552
 Db TGGCGCAGCTGTCAAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4525
 QY 1553 CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGCGCGCGCGCGCGCGCGCG 1612
 Db CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGCGCGCGCGCGCGCGCGCG 4585
 QY 1613 CGGTGTACTAGGACCGCGAGAACCGCTGGTGGCGCGAGATCCAGAGCAGGCGCGCGCGCG 1672
 Db CGGTGTACTAGGACCGCGAGAACCGCTGGTGGCGCGAGATCCAGAGCAGGCGCGCGCGCG 4645
 QY 1673 AGTGGAGCTACAGATCTACCGAGCGCTTCAAGAACCTGAAGACCGCGCGCGCGCGCGCG 1732
 Db AGTGGAGCTACAGATCTACCGAGCGCTTCAAGAACCTGAAGACCGCGCGCGCGCGCGCG 4705
 QY 1733 AGATGCGCAGCG 1792
 Db AGATGCGCAGCG 4765
 QY 1793 CCATGGAGAGCTGCTGATCTGGGCGCAAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1852
 Db CCATGGAGAGCTGCTGATCTGGGCGCAAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4825
 QY 1853 AGACCTGGGAGAGCTGCTGATCTGGGCGCAAGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 1912
 Db AGACCTGGGAGAGCTGCTGATCTGGGCGCAAGCCCGCGCGCGCGCGCGCGCGCGCGCGCG 4885
 QY 1913 TCCTGAACACCCCG 1972

Db 4886 TCGTGAACACCCCCCTTGTGAAGCTGTGTACAGCTGGAGAGGAGCCATCATCG 4945
 Qy 1973 GGGCCAGACCTTCTACGTGGAGGCGCGCCCAACCGCCAGACAGATCGGCAAGCGCG 2032
 Db 4946 GGGCCAGACCTTCTACGTGGAGGCGCGCCCAACCGCCAGACAGATCGGCAAGCGCG 5005
 Qy 2033 GCTACGTGACGACCGGCGCGGCGCAGCAAGATCGTGAGCCCTGACCGAGACCAACCAACAGA 2092
 Db 5006 GCTACGTGACGACCGGCGCGGCGCAGCAAGATCGTGAGCCCTGACCGAGACCAACCAACAGA 5065
 Qy 2093 AGACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCG 2152
 Db 5066 AGACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACAGCGGCGGTGAACATCG 5125
 Qy 2153 TGACCGACAGCAGTACGCGCTGGGCGATCATCCAGGCGCCAGCCGACAGAGCGAGCG 2212
 Db 5126 TGACCGACAGCAGTACGCGCTGGGCGATCATCCAGGCGCCAGCCGACAGAGCGAGCG 5185
 Qy 2213 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGG 2272
 Db 5186 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGGCTGGG 5245
 Qy 2273 TGCCCGCCACAGAGGCGATCGGCGGCAACAGCAGATCGACAGCTGGTGAGCAAGGGCA 2332
 Db 5246 TGCCCGCCACAGAGGCGATCGGCGGCAACAGCAGATCGAGGCTGGTGAGCGCGGCA 5305
 Qy 2333 TCCGCAAGGTGCTTCCCTGGAGCGCATCGATGGCGGCGATCGTGATCTACCACTA 2387
 Db 5306 TCCGCAAGGTGCTTCCCTGGAGCGCATCGACAAAGGCCCAAGGAGGAGAGCGAGAA 5360

RESULT 12

ID ABK91613 standard; DNA; 12411 BP.
 AC ABK91613;
 DT 14-AUG-2002 (first entry)
 DE Modified HIV protein-encoding plasmid DNA #165.
 XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
 KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
 KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
 XX Synthetic.
 OS
 XX WO200232943-A2.
 XX
 XX 25-APR-2002.
 XX
 XX 14-AUG-2001; 2001WO-US25721.
 XX
 XX 14-AUG-2000; 2000US-225097P.
 PR 14-NOV-2000; 2000US-252115P.
 PR 28-MAR-2001; 2001US-279257P.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (CHAD/) CHADRAPARTII B K.
 XX
 XX Nabel GJ, Huang Y;
 PI
 XX
 XX WPI; 2002-452382/48.
 DR
 XX
 XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
 PT its encoded protein, useful as vaccines for genetic or protein
 PT immunisation for acquired immunodeficiency syndrome or human
 PT immunodeficiency virus infection -
 XX
 XX Disclosure; Page 756-760: 794pp; English.
 PS
 XX The invention relates to a nucleic acid molecule encoding a modified HIV

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
 CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
 CC genetic or protein immunisation to a host, respectively. In particular
 CC these are useful for ameliorating the symptoms of acquired
 CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
 CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
 CC compositions are useful for treating or preventing HIV infections or
 CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
 CC plasmid DNA sequences of the invention.

XX Sequence 12411 BP; 2938 A; 3733 C; 3476 G; 2264 T; 0 other;

Query Match 81.2%; Score 2000.2; DB 24; Length 12411;
 Best Local Similarity 91.2%; Pred. No. 1e-242;

Matches 2195; Conservative 0; Mismatches 188; Indels 22; Gaps 5;

Qy 14 TGCCCGAGGCGCATGAGCCAGG---CCACAGCGCCCAACATCTCTGTATGACGCGGCAACT 70
 Db 2967 TGCCCGAGGCGCATGAGCCAGGTCACCAACAGCGCCACCATCATGATGACGCGGCAACT 3026
 Qy 71 TCRAAGGGCCCCAAGCGCATCAAGTGTCTCAACTGCGCAAGGAGGGCCACATCGCCC 130
 Db 3027 TCCGCAACAGCCCAAGATCGTGAAGTGTCTCAACTGCGCAAGGAGGGCCACACCGCCC 3086
 Qy 131 GCAACTGCCGCGCCCGCCCGCAAGAAGGGCTGTCTGGAAGTGCAGCAAGAGGGCCACACAGA 190
 Db 3087 GCAACTGCCGCGCCCGCCCGCAAGAAGGGCTGTCTGGAAGTGCAGCAAGAGGGCCACACAGA 3146
 Qy 191 TGAAGACTGACCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 250
 Db 3147 TGAAGACTGACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3205
 Qy 251 GCAAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 310
 Db 3206 GGAAGGCCAGGGAAATTTCTTACAGCAGACAGACAGACAGCCCGCCACCAAGAGAGAGC 3265
 Qy 311 TGCAGGTGCGCGG-----CGACAACCCCGCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCA 364
 Db 3266 TTCAGGTTTGGGAAGAGAGACAACCTCCCTCTCAGAAGCAGGAGCGGATAGACAGGAA 3325
 Qy 365 -----CCCTGAACCTTCCCGCAGATCACCTGTGGGAGCGCCCGCTGGTGAGCATCAAGG 418
 Db 3326 CTGTATCTTCTTAGCTTCCCTCAGATCACTCTTTGGGAGCGACCCCTGTCACTAATAAGA 3395
 Qy 419 TGGGCGCCAGATCAAGGAGGCGCTGTGTGACACCGCGCGCGCGGCGGCGGCGGCGGCGGCGG 478
 Db 3386 TAGGGGCGCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGCGGCGGCGGCGGCGGCGGCGG 3445
 Qy 479 AGATGAGCCTTGGCGCGCAAGTGGAAAGCCCAAGATGATCGGCGGCGATCGGCGGCTTCATCA 538
 Db 3446 AGATGAACCTGCCCGCGCGCTGGAAAGCCCAAGATGATCGGCGGCGATCGGCGGCTTCATCA 3505
 Qy 539 AGGTGGCCAGTAGGACCATCTGTATCGAGATCTTCGCGCAAGAGGCGCATCGGCAACCG 598
 Db 3506 AGGTGGCCAGTAGGACCATCTGTATCGAGATCTTCGCGCAAGAGGCGCATCGGCAACCG 3555
 Qy 599 TGCTGTATCGGCT 658
 Db 3566 TGCTGTATCGGCT 3625
 Qy 659 GCACCTTGAACTTCCCGCATCAGCCCCCATCGAGACCGGTGCGCGGTGAAGCTGAAGCGCGGCA 718
 Db 3626 GCACCTTGAACTTCCCGCATCAGCCCCCATCGAGACCGGTGCGCGGTGAAGCTGAAGCGCGGCA 3685
 Qy 719 TGACGCGCCCAAGGTGAACAGTGGCCCTTGACCGAGGAGAGATCAAGSCCTGACCG 778
 Db 3686 TGACGCGCCCAAGGTGAACAGTGGCCCTTGACCGAGGAGAGATCAAGSCCTGCTGG 3745
 Qy 779 CCATCTCGAGGAGATGGAGAAGGAGGCGGCAAGATCAACAAGATCGGCGCGCGGAGAACCCCT 838
 Db 3746 AGATCTGACCGAGATGGAGAAGGAGGCGGCAAGATCAACAAGATCGGCGCGCGGAGAACCCCT 3805

QY 839 ACAACACCCCGTGTTCGCCATCAAGAAAGAGAGACACCAAGTGGCGCAAGCTGGTG 898
Db 3806 AACACACCCCGTGTTCGCCATCAAGAAAGAGAGACACCAAGTGGCGCAAGCTGGTG 3805
QY 899 ACTTCCCGAGAGCTGAACAAGCGACCCAGAGACTTCTGGAGGTGACGTGGGCAATCCCCC 958
Db 3866 ACTTCCCGAGAGCTGAACAAGCGACCCAGAGACTTCTGGAGGTGACGTGGGCAATCCCCC 3925
QY 959 ACCCGCGCGCTGAAGAAGAAGAGCGTGAACGCTGCTGGAGCTGGGCGACGCCCTACT 1018
Db 3926 ACCCGCGCGCTGAACAGAGAGAGCGTGAACGCTGCTGGAGCTGGGCGACGCCCTACT 3985
QY 1019 TCAGCGTGCCTCGGAGAGAGACTTCGCAAGTACACCGCTTACCATCCCCAGCATCA 1078
Db 3986 TCAGCGTGCCTCGGAGAGAGACTTCGCAAGTACACCGCTTACCATCCCCAGCATCA 4045
QY 1079 AACAGAGACCCCGGATCGGCTACCAAGTACACGCTGCTGCCCGCAGGCTGGAGGGCA 1138
Db 4046 AACAGAGACCCCGGATCGGCTACCAAGTACACGCTGCTGCCCGCAGGCTGGAGGGCA 4105
QY 1139 GCCCGAGCATCTCCAGAGCAGATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACC 1198
Db 4106 GCCCGAGCATCTCCAGAGCAGATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGAAC 4165
QY 1199 CCGAGATCGTATCTACCA-----GGCCCCCTGTACGTGGGCGAGCAGCTGGAGATCG 1252
Db 4166 CCGAGATCGTATCTACCAAGTATGACGACGACCTGTACGTGGGCGAGCAGCTGGAGATCG 4225
QY 1253 GCCAGCACCGGCAAGATCGAGGAGCTGGCAAGCAGCTGCTGCGCTGGGGCTTCACCA 1312
Db 4226 GCCAGCACCGGCAAGATCGAGGAGCTGGCAAGCAGCTGCTGCGCTGGGGCTTCACCA 4285
QY 1313 CCCCCGACAAGACACAGAGAGCGCCCTTCTCTGTGGATGGCTACGAGCTGCACC 1372
Db 4286 CCCCCGACAAGACACAGAGAGCGCCCTTCTCTGTGGATGGCTACGAGCTGCACC 4345
QY 1373 CCGACAAGTGCACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG 1432
Db 4346 CCGACAAGTGCACCGTGCAGCCCATCGTGTGCCGAGAGAGAGCTGGACCGTGAACG 4405
QY 1433 ACATCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCAGATACCCCGGCATCAAG 1492
Db 4406 ACATCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCAGATACCCCGGCATCAAG 4465
QY 1493 TGCGCCAGCTGTGAACGCTGTGCGCGCGCAGAGCGCTGACGACATCGTGCCTGTA 1552
Db 4466 TGCGCCAGCTGTGAACGCTGTGCGCGCGCAGAGCGCTGACGAGCTGTGCGCCCTGA 4525
QY 1553 CCGAGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTCGCGAGCGCGTGCACG 1612
Db 4526 CCGAGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGAAGAGCGCGTGCACG 4585
QY 1613 GCGTGTACTAGGACCCAGCAAGACCTGGTGGCGGAGATCCAGAACGAGCGCCAGCAC 1672
Db 4586 GCGTGTACTAGGACCCAGCAAGACCTGTGCGCGGAGATCCAGAACGAGCGCGCGAGGCC 4645
QY 1673 AGTGGACCTACAGATCTACAGGAGCGCTTCAAGAACCTGAAGACCGCAAGTACGCCA 1732
Db 4646 AGTGGACCTACAGATCTACAGGAGCGCTTCAAGAACCTGAAGACCGCAAGTACGCCA 4705
QY 1733 AGATGCGACCGCGCCACCAACGACGTGAAGCAGCTGACCGAGCGCGTGCAGAAATCG 1792
Db 4706 GCATGAAGGGCGCCACCAACGACGTGAAGCAGCTGACCGAGCGCGTGCAGAAATCG 4765
QY 1793 CCATGGAGAGATCGTATCTGGGCAAGACCCCAAGTTCCGGCTGCCCATCCAGAAAG 1852
Db 4766 CCACCGAGAGATCGTATCTGGGCAAGACCCCAAGTTCAAGCTGCCCATCCAGAAAG 4825
QY 1853 AGACCTGGAGACCTGTGTGACCGACTACTGGCAGGCGCACTGTGATCCCGAGTGGGAGT 1912
Db 4826 AGACCTGGAGGCGCTGTGTGACCGAGTACTGGCAGGCGCACTGTGATCCCGAGTGGGAGT 4885
QY 1913 TCGTGAACACCCCGCTGTGTGAAGCTGTGTACCAAGCTGGAGAGAGCGCCATCATCG 1972

Db 4886 TCGTGAACACCCCGCTGTGTGAAGCTGTGTACCAAGTGGAGAGCGCCATCATCG 4945
QY 1973 GCGCGGAGAGCTTCTAGTGTGACGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCCG 2032
Db 4946 GCGCGGAGAGCTTCTAGTGTGACGCGCGCCCAACCGGAGACCAAGTGGGCAAGGCCG 5005
QY 2033 GCTACGTGACCGACCGGCGCGCAGAAAGATCGTGAAGCTGACCGAGACCAACCAACCA 2092
Db 5006 GCTACGTGACCGACCGGCGCGCAGAAAGTGTGTGCCCTGACCGACCAACCAACCA 5065
QY 2093 AGACCGAGCTGACGCGCATCCAGCTGGCCCTGCAAGGACAGCGGCGAGGTGAACATCG 2152
Db 5066 AGACCGAGCTGACGCGCATCCAGCTGGCCCTGCAAGGACAGCGGCGTGGAGTGAACATCG 5125
QY 2153 TGACCGACAGCAGTACGCGCTGGCATATCCAGGCGCGCCGACGAGAGCGGAGCG 2212
Db 5126 TGACCGACAGCAGTACGCGCTGGCATATCCAGGCGCGCCGACGAGAGCGGAGCG 5185
QY 2213 AGCTGGTGAACAGCATATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGG 2272
Db 5186 AGCTGGTGAACAGCATATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGG 5245
QY 2273 TGCCCGCCCAAGGCGCATCGCGCGCAACGAGCAGATCGACAAGCTGGTGAAGGCGCA 2332
Db 5246 TGCCCGCCCAAGGCGCATCGCGCGCAACGAGCAGGTGGAGCGGCTGGTGAAGCGCGCA 5305
QY 2333 TCCCGAAGGTGCTTCTCTGACGCGCATCGATGGGCGATCGTGTATCTACCACTA 2387
Db 5306 TCCCGAAGGTGCTTCTCTGACGCGCATCGACAAGGCGCGCAGGAGGAGCACAGAA 5360

RESULT 13

ABK91624

ID ABK91624 standard; DNA; 9785 BP.

XX ABK91624;

XX 14-AUG-2002 (first entry)

XX Modified HIV protein-encoding plasmid DNA #176.

XX HIV; human immunodeficiency virus; gene; ds; circular; anti-HIV;
XX Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
XX acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX Synthetic.

XX WO200232943-A2.

XX 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.

XX 14-AUG-2000; 2000US-225097P.

XX 14-NOV-2000; 2000US-252115P.

XX 28-MAR-2001; 2001US-279257P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX (CHAD/) CHADABARTI B K.

XX Nabel GJ, Huang Y;

XX WPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
XX its encoded protein, useful as vaccines for genetic or protein
XX immunisation for acquired immunodeficiency syndrome or human
XX immunodeficiency virus infection

XX Disclosure; Page 791-794; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV

||||| TGAACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGGAGAGGAGCCCATCATCGCGG 4937
Db
QY 1976 CCGAGACCTTCTAGTGGAGCGGCCGCCCAACCGCGAGACCAAGATCGGCAAGCCCGCT 2035
Db
QY 4938 CCGAGACCTTCTAGTGGAGCGGCCGCCCAACCGCGAGACCAAGATCGGCAAGCCCGCT 4997
QY 2036 ACGTGACCGACCGGGCGCGCAGAAATCGTGAAGCTGACCGGAGACCAACCAAGCAAGA 2095
Db
QY 4998 ACGTGACCGACCGGGCGCGCAGAAATCGTGAAGCTGACCGGAGACCAACCAAGCAAGA 5057
QY 2096 CCGAGCTGACAGGCCATCCAGCTGGCCCTGCAGGACACGCGGAGAGGTGAACATCGTGA 2155
Db
QY 5058 CCGAGCTGACAGGCCATCCAGCTGGCCCTGCAGGACACGCGCTGGAGGTGAACATCGTGA 5117
QY 2156 CCGACACAGTACCGCTGGGATCATTCAGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGG 2215
Db
QY 5118 CCGACACAGTACCGCTGGGATCATTCAGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGG 5177
QY 2216 TGGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2275
Db
QY 5178 TGGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5237
QY 2276 CCGCCCAAGAGGATCGCGGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGCGCATCC 2335
Db
QY 5238 CCGCCCAAGAGGATCGCGGCAACGAGCAGATCGACAAAGCTGGTGAGCGCGCGCATCC 5297
QY 2336 GCAAGTGTCTTCTGGAGCGGATCGATGGCGGCATCGTGATACAGTA 2387
Db
QY 5298 GCAAGTGTCTTCTGGAGCGGATCGATGGCGGCATCGATGGCGGCATCGTGATACAGTA 5349

RESULT 14

ABK91619
ID ABK91619 standard; DNA; 9167 BP.
XX
AC ABK91619;
XX
XX 14-AUG-2002 (first entry)
XX
DE Modified HIV protein-encoding plasmid DNA #171.
XX
XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX
OS Synthetic.
XX
XX WO200232943-A2.
XX
XX 25-APR-2002.
XX
XX 14-AUG-2001; 2001WO-US25721.
XX
XX 14-AUG-2000; 2000US-225097P.
XX
XX 14-NOV-2000; 2000US-252115P.
XX
XX 28-MAR-2001; 2001US-279257P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX (CHAD/) CHADABARTI B K.
XX
XX Nabel GJ, Huang Y;
XX
XX WPI; 2002-452382/48.
XX
XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
XX its encoded protein, useful as vaccines for genetic or protein
XX immunisation for acquired immunodeficiency syndrome or human
XX immunodeficiency virus infection
XX
XX Disclosure; Page 775-778; 794pp; English.
XX
XX The invention relates to a nucleic acid molecule encoding a modified HIV

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX

SQ Sequence 9167 BP; 2225 A; 2668 C; 2517 G; 1757 T; 0 other;

Query Match 80.9%; Score 1993.4; DB 24; Length 9167;
Best Local Similarity 91.1%; Pred. No. 7.6e-242;
Matches 2182; Conservative 0; Mismatches 186; Indels 27; Gaps 5;
QY 14 TGGCCGAGGGCCATGAGCCAGG---CCACCAGCGCCCAACATCCTGATGAGCGCAGCAACT 70
Db 2964 TGGCCGAGGGCCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGAGCGCGCAACT 3023
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTCTTCAACTGGGGCAAGAGGGCCACATTCGCCC 130
Db 3024 TCCGCAACACGCGCAAGATCGTGAAGTCTTCAACTTGGCGAAGAGGGCCACACCGCCC 3083
QY 131 GCAACTCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGCGAAGAGGGCCACACAGA 190
Db 3084 GCAACTCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGCGAAGAGGGCCACACAGA 3143
QY 191 TGAAGGACTGCACCGAGCGCCAGCGCCCAACTTCTTCCGCGAGGAGCTGGCCCTTCCCCCAGG 250
Db 3144 TGAAGGACTGCACCGAGCGCACAGGCTAA-----TAGGGAAGATCTGCGCTTCCACAAAG 3197
QY 251 GCAAGGGCCCGGAGTTTCCCGAGCGAGAGAACCGCGGCCCAACAGCCCAACAGCGCGCAGC 310
Db 3198 GGAAGGGCAGGGAATTTTCTTCAAGAGCAGACAGAGCCCAACAGCGCCCAACAGAGAGC 3257
QY 311 TGCAGGTGCGGG-----CGACAACCCCGCGAGGCGCGCGCGCGCGCGCGCGCGCA 364
Db 3258 TGCAGGTGCGGGAGAGAGAACAACTCCCTCTCAGAAAGAGGAGCGCGATAGCAAGGAA 3317
QY 365 -----CCCTGAACTTCCCGAGATCACTGTGGAGCGCCCTGTGTGAGCATCAAG 418
Db 3318 CTGTATCTTTAGCTTCCCTCAGATCACTCTTGGCAGCGACCCCTCGTCAATAAAGA 3377
QY 419 TGGCGGCGAGATCAAGAGGCGCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGCGAGG 478
Db 3378 TAGGGGCGAGCTGAAGGAGGCGCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGAGG 3437
QY 479 AGATGAGCTGCCCGCAAGTGAAGCCCAAGATGATCGCGCGGATCGCGCGCTTCATCA 538
Db 3438 AGATGAACCTGCCCGCGCGCTGGAAGCCCAAGATGATCGCGCGGATCGCGCGCTTCATCA 3497
QY 539 AGGTGCGCGAGTACGACAGATCTCTGATCGAGATCTGCGGCAAGAGGCGCATCGCACCG 598
Db 3498 AGGTGCGCGAGTACGACAGATCTCTGATCGAGATCTGCGGCGCAAGGCGCATCGCACCG 3557
QY 599 TGCTGATCGGCGCCACCGCGCTGAACATCATCGCGCGCAACATGCTGACCGAGCTGGGCT 658
Db 3558 TGCTGATCGGCGCCACCGCGCTGAACATCATCGCGCGCAACATGCTGACCGAGATCGGCT 3617
QY 659 GCACCTGTAACCTTCCCATCATGAGCCCATCGAGACCGTGCCTGAAGTGAAGCGCGCA 718
Db 3618 GCACCTGTAACCTTCCCATCATGAGCCCATCGAGACCGTGCCTGAAGTGAAGCGCGCA 3677
QY 719 TGGAGCGGCCCAAGTGAAGCAGTGGCCCTTCAACCGAGGAGAGATCAAGCCCTTGACCG 778
Db 3678 TGGAGCGGCCCAAGTGAAGCAGTGGCCCTTCAACCGAGGAGAGATCAAGCCCTTGACCG 3737
QY 779 CCATCTCGGAGGAGATGGAGAGGAGGCAAGATCAACAGATCGGCGCGCGCGCGCGAGACCCCT 838
Db 3738 AGATCTGACCGAGATGGAGAGGAGGCAAGATCAACAGATCGGCGCGCGCGCGCGAGACCCCT 3797

QY 839 ACAACACCCCGGTTTCGCCATCATAGAAAGAGACACACCAAGTGGCGGAAGCTGTGG 898
Db 3798 ACAACACCCCGGTTTCGCCATCAAGAAAGAGACACCAAGTGGCGGAAGCTGTGG 3857
QY 899 ACTTCCCGAGCTGAACAAGCAGCAGGACTTCTGGGAGTGCAGCTGGGCATCCCC 958
Db 3858 ACTTCCCGAGCTGAACAAGCAGCAGCAGGACTTCTGGGAGTGCAGCTGGGCATCCCC 3917
QY 959 ACCCGCGCGCTGAAGAAGAAAGAGCGTGACCGTGTGGAGCTGGGCGACGCCCTACT 1018
Db 3918 ACCCGCGCGCTGAACAGAGAGAGCGTGACCGTGTGGAGCTGGGCGACGCCCTACT 3977
QY 1019 TCAGCTGCCCTGGAGCAGGACTTCCGCAAGTACACCGCCTTCAACATCCCAAGCATCA 1078
Db 3978 TCAGCTGCCCTGGAGCAGGACTTCCGCAAGTACACCGCCTTCAACATCCCAAGCATCA 4037
QY 1079 ACAACAGACCCCGGATCCGCTACCAAGTACACAGTGTGCCCCAGGCGTGGAGGGCA 1138
Db 4038 ACAACAGACCCCGGATCCGCTACCAAGTACACAGTGTGCCCCAGGCGTGGAGGGCA 4097
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATTCCTGGAGCCTTCCGGCGCCGCAACC 1198
Db 4098 GCGCCAGCATCTTCCAGTGCAGCATGACCAAGATTCCTGGAGCCTTCCGGCAAGCAGAAC 4157
QY 1199 CCGAGATCGTATCTACCA-----GCCCGCCTGTACGTGGGCGAGCAGCTGGAGATCG 1252
Db 4158 CCGACATCGTATCTACAGTATGACCAACCTGTACGTGGGCGAGCAGCTGGAGATCG 4217
QY 1253 GCAGACCCGCGCAAGTGCAGGAGCTGCAGAGCAGCTGCTGCGCTGGGCTTCACCA 1312
Db 4218 GCAGACCCGCGCAAGTGCAGGAGCTGCAGAGCAGCTGCTGCGCTGGGCTTCACCA 4277
QY 1313 CCGCCGCAAGAGCAGCAG 1372
Db 4278 CCGCCGCAAGAGCAGCAG 4337
QY 1373 CCGACAGTGGAGCTGCAGCCCATCGAGCTGCCGAGAGAGAGAGAGAGAGAGAGAGAG 1432
Db 4338 CCGACAGTGGAGCTGCAGCCCATCGAGCTGCCGAGAGAGAGAGAGAGAGAGAGAGAG 4397
QY 1433 ACATCCAGAGAGCTGGTGGGCAAGTGAACCTGGCCAGCAGAGTCTACCCGCGCATCAAG 1492
Db 4398 ACATCCAGAGAGCTGGTGGGCAAGTGAACCTGGCCAGCAGAGTCTACCCGCGCATCAAG 4457
QY 1493 TGGCCAGCTGTCAAGCTCTCGCGGCGCCAAAGGCCCTGACCCAGCATCGTCCCGCTGA 1552
Db 4458 TGGCCAGCTGTCAAGCTCTCGCGGCGCCAAAGGCCCTGACCCAGCATCGTCCCGCTGA 4517
QY 1553 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGCGGAGCCGTCGACG 1612
Db 4518 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGAGAGCCGCTGCAAG 4577
QY 1613 GCCTGTACTACGACCCCGAG 1672
Db 4578 GCCTGTACTACGACCCCGAG 4637
QY 1673 AGTGGACCTACAGATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1732
Db 4638 AGTGGACCTACAGATCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4697
QY 1733 AGATGGCAGCCGCGCAGCAGCAGTGAAGCAGTGCAGGAGCGGTCGAGAGATCG 1792
Db 4698 GCATGAAGGGCGCCACACCAAGCAGTGAAGCAGTGCAGGAGCGGTCGAGAGATCG 4757
QY 1793 CCATGGAGAGCATCTGATCTGGGGAAGACCCCAAGTTCCCGCTGCCCATCCAGAGG 1852
Db 4758 CCACCGAGAGCATCTGATCTGGGGAAGACCCCAAGTTCAAGCTGCCCATCCAGAGG 4817
QY 1853 AGACCTGGGAGACCTGGTGAGCAGTACTTGGCAGGCGACCTGGATCCCGAGTGGGAGT 1912
Db 4818 AGACCTGGGAGGCTGGTGAGCAGTACTTGGCAGGCGACCTGGATCCCGAGTGGGAGT 4877
QY 1913 TCGTGAACACCCCGGTTGAGCTGTGGTACCAAGTGGAGAGAGAGAGAGAGAGAGAGAG 1972

Db 4878 TCGTGAACACCCCGGTTGAGAGCTGTGGTACCAAGTGGAGAGAGAGAGAGAGAGAGAGAG 4937
QY 1973 GCGCCGAGAGACTTCTAGCTGGAGCGCGCCCAACCGGAGAGAGAGAGAGAGAGAGAGAG 2032
Db 4938 GCGCCGAGAGACTTCTAGCTGGAGCGCGCGCCCAACCGGAGAGAGAGAGAGAGAGAGAG 4997
QY 2033 GCTACGTGACCGAGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2092
Db 4998 GCTACGTGACCGAGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5057
QY 2093 AGACCGAGCTGCAGGCGCATCAGCTGGCCCTGTCAGGAGAGAGAGAGAGAGAGAGAGAG 2152
Db 5058 AGACCGAGCTGCAGGCGCATCAGCTGGCCCTGTCAGGAGAGAGAGAGAGAGAGAGAGAG 5117
QY 2153 TGACCGAGAGCGAGTACGCGCTGGGCATCATCCAGGCGCGAGCGCGAGAGAGAGAGAGAG 2212
Db 5118 TGACCGAGAGCGAGTACGCGCTGGGCATCATCCAGGCGCGAGCGCGAGAGAGAGAGAGAG 5177
QY 2213 AGCTGTGTAACAGAGATCATCGAGCAGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAG 2272
Db 5178 AGCTGTGTAACAGAGATCATCGAGCAGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAG 5237
QY 2273 TGCCCGCGCCACAGAGGAGTGGCGGCAACGAGCAGATCGGAGAGAGAGAGAGAGAGAGAG 2332
Db 5238 TGCCCGCGCCACAGAGGAGTGGCGGCAACGAGCAGAGTGGAGCGCTGTGAGCGCGCGCA 5297
QY 2333 TCAGCAGAGTGTCTCTCTGAGCGGAGTGCATGCGCGCATCTGTATCCAGTA 2387
Db 5298 TCAGCAGAGTGTCTCTCTGAGCGGAGTGCATGAGAGCGCGAGAGAGAGAGAGAGAGAG 5352

RESULT 15

ABK91614

ID ABK91614 standard; DNA; 9170 BP.

XX ABK91614;

XX 14-AUG-2002 (first entry)

XX Modified HIV protein-encoding plasmid DNA #166.

XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
XX Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
XX acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX Synthetic.

XX WO200232943-A2.

XX 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.

XX 14-AUG-2000; 2000US-225097P.

XX 14-NOV-2000; 2000US-252115P.

XX 28-MAR-2001; 2001US-279257P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX (CHAD/) CHADABARTI B K.

XX Nabel GJ, Huang Y;

XX WPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
XX its encoded protein, useful as vaccines for genetic or protein
XX immunisation for acquired immunodeficiency syndrome or human
XX immunodeficiency virus infection -

XX Disclosure; Page 760-763; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

XX
SQ Sequence 9170 BP; 2225 A; 2669 C; 2519 G; 1757 T; 0 other;

Query Match 80.9%; Score 1993.4; DB 24; Length 9170;
Best Local Similarity 91.1%; Pred. No. 7.6e-242;
Matches 2182; Conservative 0; Mismatches 186; Indels 27; Gaps 5;

QY 14 TGGCCGAGGCATCAGCCAGG---CCACCAGCGCCACATCTCTGATCAGCGGAGCAACT 70
DB TGGCCGAGGCATCAGCCAGGTCAGTACACAGCGCCACCATCATGATCAGCGCGGCAACT 3026

QY 71 TCAAGGGCCCAACGCGCATCATCAAGTCTTCAACTGCGCAAGGAGGGCCACATCGGCC 130
DB TCAAGGGCCCAACGCGCATCTGTGAAGTCTTCAACTGCGCAAGGAGGGCCACACCGGCC 3086

QY 131 GCAACTCGCGGCCCCCGCGCAAGAGGCTGCTGGAAGTGGCGGAAGGAGGGCCACACAGA 190
DB GCAACTCGCGGCCCCCGCGCAAGAGGCTGCTGGAAGTGGCGGAAGGAGGGCCACACAGA 3146

QY 191 TGAAGGACTGCACGAGCGCCAGGCCACTTCTCCGCGAGGAGCTGGCTTCCCCCAGG 250
DB TGAAGGACTGCACGAGCGCCAGGCGCTAA-----TAGGGAGATCTGGGCTTCCCACAAG 3200

QY 251 GCAAGGCGCGGAGTTTCCCGAGCGAGCAGAGAACCGCGCCACAGCCACACAGCGCGAGC 310
DB GCAAGGCGCGGAGTTTCTCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3260

QY 311 TGCAGTTCGCGGG-----CGACAAACCGCGCAGCGAGCGCGCGCGCGCGAGCGCGAGGCA 364
DB TGCAGTTCGCGGGAGAGACAAACAACTCCCTCTCAGAGCAGGAGGAGCGCGATAGACAAAGAA 3320

QY 365 -----CCCTGAACCTTCCCGAGATACCTCTGTGGCAGCGCCCGCCCTGGTGAACATCAGG 418
DB CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3380

QY 419 TGGCGGCGCAGATCAAGGAGGCGCTGTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCGG 478
DB TGGCGGCGCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGG 3440

QY 479 AGATGAGCTTCCCGGCAAGTGGAAAGCCCAAGATGATCGGGGGCATCGGGGGCTTCATCA 538
DB AGATGAACTTCCCGGCGCGCTGGAAGCCCAAGATGATCGGGGGCATCGGGGGCTTCATCA 3500

QY 539 AGTGGCCAGTAGACACAGATCTGTGAGATCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 598
DB AGTGGCCAGTAGACACAGATCTGTGAGATCTCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 3560

QY 599 TGTGATCGCGCCCGCCCGCGTGAACATCATCGCGCCGCAACATGCTGACCGCGAGTGGGCT 658
DB TGTGATCGCGCCCGCCCGCGTGAACATCATCGCGCCGCAACATGCTGACCGCGAGTGGGCT 3620

QY 659 GCACCTGAATCTTCCCATCAGCGCCCATCGAGACCGTGGCCCGTGAAGCTGAAGCGCGCGCA 718
DB GCACCTGAATCTTCCCATCAGCGCCCATCGAGACCGTGGCCCGTGAAGCTGAAGCGCGCGCA 3680

QY 719 TGGAGCGCCCAAGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
DB TGGAGCGCCCAAGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 3740

QY 779 CCATCTCGAGGAGATGGAGAGGAGGCGCAAGATCAACAGATCGGCGCCCGCGAGAACCCCT 838
DB CCATCTCGAGGAGATGGAGAGGAGGCGCAAGATCAACAGATCGGCGCCCGCGAGAACCCCT 3800

QY 3741 AGATCTGCACCGAGATGGAGAGGAGGCGCAAGATCAACAGATCGGCGCCCGCGAGAACCCCT 838
DB AGATCTGCACCGAGATGGAGAGGAGGCGCAAGATCAACAGATCGGCGCCCGCGAGAACCCCT 3800

QY 839 ACAACACCCCGTGTTCGCCATCAAGAAAGAGCAGACCAACCAAGTGGCGCAAGCTGGTGG 898
DB ACAACACCCCGTGTTCGCCATCAAGAAAGAGCAGACCAACCAAGTGGCGCAAGCTGGTGG 3860

QY 899 ACTTCCGCGAGCTGAACAAGCGCAGCCAGGACTTCTGGAGGTGAGCTGGGCGATCCCC 958
DB ACTTCCGCGAGCTGAACAAGCGCAGCCAGGACTTCTGGAGGTGAGCTGGGCGATCCCC 3920

QY 959 ACCCGCGCGGCTGAAGAAGAGAGCGTGCACCGTCTGGACCTGGCGACGCGCTACT 1018
DB ACCCGCGCGGCTGAAGAAGAGAGCGTGCACCGTCTGGACCTGGCGACGCGCTACT 3980

QY 1019 TCAGCGTCCCGCTGGACGAGGACTTCCGCAAGTACACCGCCCTTACCATCCCCAGCATCA 1078
DB TCAGCGTCCCGCTGGACAAGGACTTCCGCAAGTACACCGCCCTTACCATCCCCAGCATCA 4040

QY 1079 ACAAGGAGACCCCGGCATCCGCTACCAAGTACACAGTCTGCTCCCGCAGGCTGGGAAGGCA 1138
DB ACAAGGAGACCCCGGCATCCGCTACCAAGTACACAGTCTGCTCCCGCAGGCTGGGAAGGCA 4100

QY 1139 GCGCCAGCATCTCCAGAGCAGCATGACCAAGATCCTTGAGCCCTTCCGCGCCCGCAACC 1198
DB GCGCCAGCATCTCCAGAGCAGCATGACCAAGATCCTTGAGCCCTTCCGCGCCCGCAACC 4160

QY 1199 CCGAGATCGTGTATACCA-----GGCCCCCTGTACGTGGGCGAGCAGCTGGAGATCG 1252
DB CCGAGATCGTGTATACCAAGTACATGAGCAGCAGTGTACGTGGGCGAGCAGCTGGAGATCG 4220

QY 1253 GCCAGCAGCGCGCAAGATCAGAGGAGTGGCGAAGCAGCCTGCTGGGCTGGGGCTTCACCA 1312
DB GCCAGCAGCGCGCAAGATCAGAGGAGTGGCGAAGCAGCCTGCTGGGCTGGGGCTTCACCA 4280

QY 1313 CCCCCGCAAGAAGCACCAGAAAGGAGGCCCTTCTGTGGATGGGTACGAGCTGCACC 1372
DB CCCCCGCAAGAAGCACCAGAAAGGAGGCCCTTCTGTGGATGGGTACGAGCTGCACC 4340

QY 1373 CCGAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG 1432
DB CCGAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG 4400

QY 1433 ACATCCAGAACTGGTGGGCAAGCTGAGTGGGCGCAGCAGATCTACCCCGGCATCAAG 1492
DB ACATCCAGAACTGGTGGGCAAGCTGAGTGGGCGCAGCAGATCTACCCCGGCATCAAG 4460

QY 1493 TGGCGCAGCTGTCAAGCTGTGCGCGCGCAGAGCCCTGACGACATCTGTCGCCCTGA 1552
DB TGGCGCAGCTGTCAAGCTGTGCGCGCGCAGAGCCCTGACGAGGTGGTGGCCCTTGA 4520

QY 1553 CCGAGAGGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGCGGAGCGCCGTGCACG 1612
DB CCGAGAGGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGCGGAGCGCCGTGCACG 4580

QY 1613 GCGTGTACTAGACCCCGCAGCAAGACCTGTGGCGGAGATCCAGAGCAGGCGCCAGACC 1672
DB GCGTGTACTAGACCCCGCAGCAAGACCTGTGGCGGAGATCCAGAGCAGGCGCCAGACC 4640

QY 1673 AGTGGACCTACCAAGTCTACCAAGGAGCCCTTCAAGAACTGAAAGACCGGCAAGTACGCA 1732
DB AGTGGACCTACCAAGTCTACCAAGGAGCCCTTCAAGAACTGAAAGACCGGCAAGTACGCA 4700

QY 1733 AGATGCGACCGCCCGCACCAACGAGCTGAGCAGCTGACCGAGGCGCGTGCAGAGATCG 1792
DB AGATGCGACCGCCCGCACCAACGAGCTGAGCAGCTGACCGAGGCGCGTGCAGAGATCG 4760

QY 1793 CCATGGAGAGCATCGTGTGGGGCAAGACCCCAAGTTCGGCTGCCCATCCAGAGG 1852
DB CCATGGAGAGCATCGTGTGGGGCAAGACCCCAAGTTCGGCTGCCCATCCAGAGG 4820

QY 1853 AGACTGGGAGCCTGTGGAGCCGACTTACTGGCAGGCGCCACCTGGATCCCCAGTGGAGT 1912
DB AGACTGGGAGCCTGTGGAGCCGACTTACTGGCAGGCGCCACCTGGATCCCCAGTGGAGT 4880

QY 1913 TCGTGAACACCCCGCTGTGTAAGCTGTGGTACCAAGCTGGAGAGGAGGCCCATCATCG 1972

```
|||||
Db 4881 TCGTGAACACCCCCCTGTTGAAGCTGTGTACAGCTGGAGAAAGGAGCCCATCATCG 4940
QY 1973 GCGCCGAGACCTTCTACGTGGAGGGGCCGCCAACCGCGAGACCAAGATCGGCAAGGCCG 2032
Db 4941 GCGCCGAGACCTTCTACGTGGAGGGGCCGCCAACCGCGAGACCAAGCTGGGCAAGGCCG 5000
QY 2033 GCTACGTGACCGACCGGGGCCGAGAAGATCGTGAAGCTGACCGAGACCACCAACACAGA 2092
Db 5001 GCTACGTGACCGACCGGGGCCGAGAGGTGTGCCCTGACCGACACCAACCAACAGA 5060
QY 2093 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACACGCGGACAGGTTGAACATCG 2152
Db 5061 AGACCGAGCTGCAGGCCATCCACCTGGCCCTGCAGGACACGCGCCTGGAGGTGAACATCG 5120
QY 2153 TGACCGACACCCAGTAGCCCTGGGCATCATCCAGGCCACGCGGACAGAGCGAGAGCG 2212
Db 5121 TGACCGACACCCAGTAGCCCTGGGCATCATCCAGGCCACGCGCCTGAGAGGCGAGAGCG 5180
QY 2213 AGCTGGTGAACACAGATCATCGACAGCTGATCAAGAGGAGAAAGGTGTACCTGAGCTGGG 2272
Db 5181 AGCTGGTGAACACAGATCATCGACAGCTGATCAAGAGGAGAAAGGTGTACCTGGCCTGGG 5240
QY 2273 TGCCCGCCCAACAAGGGCATCGCGGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGGCA 2332
Db 5241 TGCCCGCCCAACAAGGGCATCGCGGCAACGAGCAGGTGGAGCGCCTGGTGAGCGCGCGCA 5300
QY 2333 TCCGCAAGGTGCTGTTCTTCTGGAGGCCATCGATGGCGGCATCGTGTATCTACCAGTA 2387
Db 5301 TCCGCAAGGTGCTGTTCTTCTGGAGGCCATCGACAAAGGCCACGAGGAGGACACGAGAA 5355
```

Search completed: February 10, 2003, 12:44:09
Job time : 583 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 11:44:48 ; Search time 3386 Seconds
(without alignments)
11780.710 Million cell updates/sec

Title: US-09-610-313-31

Perfect score: 2463

Sequence: 1 gtcgacgcaccatggccga.....gggctagcaccgggaattc 2463

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:*
1:  em_estba:*
2:  em_estmb:*
3:  em_estin:*
4:  em_estnu:*
5:  em_estov:*
6:  em_estpl:*
7:  em_estro:*
8:  em_hic:*
9:  gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86.8	3.5	2598	11 AY103647	AY103647 Zea mays
2	85.6	3.5	951	13 BM321451	BM321451 rockefell
3	80.2	3.3	1132	13 BM320864	BM320864 rockefell
4	79.6	3.2	1165	13 BM320900	BM320900 rockefell
5	79	3.2	867	13 BM321430	BM321430 rockefell
6	76.6	3.1	1550	13 BM321022	BM321022 rockefell

7	75	3.0	862	13	BM321023	rockefell
8	73.2	3.0	853	13	BM321393	rockefell
9	72.8	3.0	566	13	BM587428	BM587428 170006873
10	72.2	2.9	640	10	BE601575	BE601575 HVSME009
11	71.6	2.9	500	13	BM372120	BM372120 EBR003_SQ
12	71.6	2.9	538	13	BM368580	BM368580 EBR008_SQ
13	71.6	2.9	540	9	AJ471121	AJ471121 AJ471121
14	71.6	2.9	566	14	BQ464692	BQ464692 HF02P20r
15	71.6	2.9	578	13	BM377112	BM377112 EBR005_SQ
16	71.6	2.9	579	12	BF253640	BF253640 HVSMEF000
17	71.6	2.9	582	12	BM372871	BM372871 EBR004_SQ
18	71.6	2.9	588	14	BQ765498	BQ765498 EBR003_SQ
19	71.6	2.9	600	9	AL508778	AL508778 AL508778
20	71.6	2.9	625	14	BQ768295	BQ768295 EBR008_SQ
21	71.6	2.9	656	13	BM370222	BM370222 EBR009_SQ
22	71.6	2.9	660	14	BQ762385	BQ762385 EBR001_SQ
23	71.6	2.9	700	9	AL508513	AL508513 AL508513
24	71.6	2.9	768	12	BG365317	BG365317 HVSME1000
25	71.6	2.9	789	12	BF626988	BF626988 HVSMEB000
26	71.6	2.9	822	12	BG417596	BG417596 HVSMEF001
27	71.4	2.9	730	12	BF260696	BF260696 HVSMEF002
28	71.2	2.9	646	13	BM645449	BM645449 170006873
29	71.2	2.9	724	13	BM584191	BM584191 170006872
30	71.2	2.9	1460	11	AY104825	AY104825 Zea mays
31	71.2	2.9	2129	11	AY109914	AY109914 Zea mays
32	71	2.9	925	17	CNS0091P	AL053013 Drosophila
33	70.8	2.9	2299	11	AY106831	AY106831 Zea mays
34	70.6	2.9	861	10	BE636696	BE636696 rockefell
35	70.2	2.9	3134	11	AY109500	AY109500 Zea mays
36	70	2.8	494	12	BG349122	BG349122 947029E04
37	70	2.8	571	12	BF266979	BF266979 HV_CEA001
38	70	2.8	609	13	BJ292939	BJ292939 BJ292939
39	70	2.8	613	13	BJ226790	BJ226790 BJ226790
40	70	2.8	616	13	BJ280623	BJ280623 BJ280623
41	70	2.8	619	13	BJ291714	BJ291714 BJ291714
42	70	2.8	627	12	BG300932	BG300932 HVSMEB001
43	70	2.8	635	13	BJ290451	BJ290451 BJ290451
44	70	2.8	643	13	BJ289971	BJ289971 BJ289971
45	70	2.8	652	13	BJ290578	BJ290578 BJ290578

ALIGNMENTS

RESULT 1
AY103647
LOCUS AY103647
DEFINITION Zea mays PC0142084 mRNA sequence.
ACCESSION AY103647
VERSION AY103647.1 GI:21206725
KEYWORDS HTPC.
SOURCE Zea mays.
ORGANISM Zea mays

REFERENCE 1 (bases 1 to 2598)
Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
Unpublished (2002)
REFERENCE 2 (bases 1 to 2598)
Coe, E.C.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA

FEATURES
source
/organism="Zea mays"
/db_xref="MaizeDB:638378"
/db_xref="taxon:4577"
/clone="PC0142084"


```

RESULT 2
BM321451          951 bp mRNA linear EST 03-JAN-2002
LOCUS
DEFINITION
rockefeller.0.1211 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
3.3.1.1), mRNA sequence.
ACCESSION
BM321451          GI:18055857
KEYWORDS
EST.
SOURCE
Mastigamoeba balamuthi.
ORGANISM
Mastigamoeba balamuthi.
REFERENCE
1 (bases 1 to 951)
AUTHORS
Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.
TITLE
The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE
21819461
COMMENT
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 951 Std Error: 0.00
POLYA-No.
FEATURES
source
1. .951
/organism="Mastigamoeba balamuthi"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"
BASE COUNT 186 a 321 c 303 g 139 t 2 others
ORIGIN
Query Match 3.5%; Score 85.6; DB 13; Length 951;
Best Local Similarity 46.0%; Pred. No. 0.0001;
Matches 323; Conservative 0; Mismatches 376; Indels 3; Gaps 1;

QY 1728 CGCAAGATCGGCACCCACACACGAGCGTGAAGCAGTGCACCGAGCGCGTGCAGAA 1787
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 CGCAAGATCGGTGGTCTGTCACATCTTCTCGACGAGGACACCGCGCGCGC 278
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1788 GATCGCCATGGAGAGCATCGTATCTGGGCGACAGCCCAAGTTCCGCTGCCATCCA 1847
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 279 CATCGCGCAGCGCGGCTCTCGGTCTTCGCTTGAAGGGGGAACCTCCAGGAGTACTG 338
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1848 GAAGGAGACCTGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTG 1907
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 339 GGAGTGACCTGGAGGCCCTGTCTTCGGCCCTACGAGGCCCTCAGATCATCTCGA 398
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1908 GGAGTTCTGTGAACACCCGCCCTGGTGAAGCTGTGTACGAGTGGAGAGGAGCCCAT 1967
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 399 CGAGCGGGTGACGGGACTCTGATGATCCACAGGGGTTCGCGCCGAGGACACCCCAA 458
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1968 CATCGGGCGCGACCTTCTACGTGGACGGCGCGCCCAACCGGAGACCAAGATCGGCAA 2027
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 459 GCTGCTGGAGGACACGAGGGCCTCGAGGAGGTGCGCTCTCAACACGCTGCTCAAGCA 518
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 2028 GGCGGGCTAGTGACCAACCGGGCGCGGACAGATGCTGAGCCTGACCGAGACCCAA 2087
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 519 GGTCAGAGGAGAGCGCGGCTTCTGGCACAGATCCTCCCGAGATCCGCGGTGTCAG 578
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 2088 CCAGAACCGAGCTGCAGGCAATCCAGCTGGCCCTGCAGACAGCGAGCGAGGTGAA 2147
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 579 CGAGGACGACGACTGGCGTATGAGGCTGTACCACTGCACCGCGACGCAAGCTGCT 638
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 2148 CATGCTGACGACGACGAGTACGCCCTTGGGATCATCCAGGCCACCGCGCAAGAGCGA 2207
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 639 GTTCCCGCGCGCTCAACGCTCAACGAC---TCTNTCACCACGAGCAAGTTCGACACATCTA 695
QY 2208 GAGGAGCTGTTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTTGAG 2267
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 696 CGGTGCGGCACCTCGCTCATCGACGGCANTCAACGGCGGACCCAGCTGATGCTCGGCGG 755
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2268 CTGGTGCCCGCCACAAAGGGCATCTCGCGGCACAGCAGCATCGACAAAGCTTGGTGAGCAA 2327
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 756 CAAGTCTGCGCTGCTGCGCGGCTACGCGGCGAGCTGGGCAAGGGCTGCGCGAGTGCCTGCG 815
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2328 GGGCATCCGCAAGGTGCTGTTCTTGAGCGGATCGATGGCGGATCGTGATCTACCAAGTA 2387
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 816 CGCGCAGGGCTGCGCGCTCATCTGACGAGATCGACCCCATCTGCGCGCTGCAGGCGTC 875
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2388 CATGGAGACCTGTACCTGGGCGAGCGCGGCCCTAGGATCGA 2429
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 876 GATGGCGCGCTTCAGGTCAACACGCTCGAGGCGGGGCTCGA 917
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
BM320864          1132 bp mRNA linear EST 03-JAN-2002
LOCUS
DEFINITION
rockefeller.0.46 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
sequence.
ACCESSION
BM320864          GI:18055270
KEYWORDS
EST.
SOURCE
Mastigamoeba balamuthi.
ORGANISM
Mastigamoeba balamuthi.
REFERENCE
1 (bases 1 to 1132)
AUTHORS
Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.
TITLE
The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE
21819461
COMMENT
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 1132 Std Error: 0.00
POLYA-No.
FEATURES
source
1. .1132
/organism="Mastigamoeba balamuthi"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"
BASE COUNT 220 a 413 c 324 g 171 t 4 others
ORIGIN
Query Match 3.3%; Score 80.2; DB 13; Length 1132;
Best Local Similarity 45.3%; Pred. No. 0.0008;
Matches 375; Conservative 0; Mismatches 444; Indels 9; Gaps 2;

QY 139 CGCGCCCCCGCAAGAGGGCTGCTGGAAGTTCGGCAAGAGGGGCCACAGATGAAGGAC 198
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 CGCGCGAGGSCAAGACGGACTACCGCGCGCCANCTGGTGATCCAGGACAAGAACAG 138
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 199 TGCACCGAGCGCCAGGCCAATCTTTCGCGAGGACCTTGGCCCTTCCCGCCAGGCGCC 258
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 TACAAACGCCCAAGTACCGCTTGTCTGTC-----GTCCGCTTTCACCAACAGGACATCGTC 192
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 259 CGCGAGTTCGCCACGAGCAGAACCGGCCAACAGCCCCACCGCGCGAGCTGCAGGTG 318
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 TGCAGATCGCCTACGCCAAGATCGACGCGGACCAATCTCTCGCGCGCCCTACTCGCAC 252
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 319 CGCGCGACAAACCCCGCAGCGAGGCGCGCGCCGAGCGCCGAGGACCCCTGAACCTTCCCC 378
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 253 GAGCTACCCGCTTCGGCGTCAAGCTCGGCTGACCAACTAGCGCGCGCTACGCGACT 312
Qy 379 CAGATACCTGTGGCAGCGCCCTGGTGAGCATCAAGTGGCGGCGAGATCAAGGAG 438
Db 313 GGCTGTGTGGCGCCCGCTGTGTAAGAAGCTCAACCTCGACTCCAAGTACGAGGT 372
Qy 439 GCGCTGTGGACACCGCGCGGAGACACGCTGTGAGGAGATGAGCCTGCCGCGCAAG 498
Db 373 GTCAAGAGGTCAACGGGAGGACTACAGTTCGAGAGTTCGACGACGGCGCCGCGGT 432
Qy 499 TGAAGCCCAAGATGATCGCGGCGCATCGGCGGCTTCAAGTGGCGGCGAGTACGACAG 558
Db 433 TTCAAGCCCTGTCTGACGCTGGCGCTGTGCGCACCTCGACTGCGCGCGGCTTCCGCC 492
Qy 559 ATCCTGATCAGATCTCGGCGCAGAAGGCCATCGGCACCGTGTGATCGGCGCCACCC 618
Db 493 GCGCTCAAGGGCATGTGCGAGCGGCGGTCAAGTCCCC---CACAGCGAGACCGGCTTC 549
Qy 619 GTGAACATCATCGCGCCCAACATGCTGACCCAGTGGGCTGCACCTGAACCTCCCCATC 678
Db 550 GTCGGCTTCAACGGCGACAGAAGGAGCTCAACCGCGGTCTCTCGCAAGTACATCTTC 609
Qy 679 AGCCCCATCAGAGCCGTGCGCGTGAAGCTGAAGCCCGGCATGGACGGCCGCCAAGGTGAAG 738
Db 610 GCGCGCCACGCTCGCGCGGTACATGAAGCTCTCAAGGAGCAGGAGCGCGCGCTTCGAC 669
Qy 739 CAGTGGCCCTGACCGAGGAGAGATCAAGCCCTGACCCCATCTCGCGAGGAGATGGAG 798
Db 670 CGCGAGTTCGCGCTACGCCAAGGAGGTGTACCGCGGACATGCTCGAGAAGATCTAC 729
Qy 799 AAGGAGGCAAGATCAACCAAGATCGCGCGGAGAACCCCTACACACCCCGCGTTCGCC 858
Db 730 ACCGAGCCACAGAGATCGCGCGGACCCGACCTTCGTCCCAAGCGCGCTCGAAG 789
Qy 859 ATCAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCGCGGAGCTGAACAAG 918
Db 790 CCGAGGGCCCAAGCCCAAGCACTGGGCAAGCGGAGGTGACGTACACAGCGCGCAAG 849
Qy 919 CGCACCAGGACTTCTGGGAGTGCAGCTGGGATCCCCACCCCGCC 966
Db 850 AACCGGTGCGCCAGAGAGGTCCGCTGGGCTACCCCGACGCGCC 897

```

RESULT 4
BM320900
LOCUS
DEFINITION
rockefeller.0.353 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
sequence.
BM320900
BM320900.1 GI:18055306
EST.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
21819461

Contact: Muller Miklos

Laboratory of Biochemical Parasitology

The Rockefeller University

1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockefeller.edu

Insert Length: 1165 Std Error: 0.00

POLYA-No. Location/Qualifiers

FEATURES

source

1..1165

Query Match 3.2%; Score 79.6; DB 13; Length 1165;
Best Local Similarity 45.1%; Pred. No. 0.001;
Matches 428; Conservative 0; Mismatches 506; Indels 16; Gaps 3;
BASE COUNT 221 a 426 c 336 g 178 t 4 others
ORIGIN

/organism="Mastigamoeba balamuthi"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"
221 a 426 c 336 g 178 t 4 others

Qy 90 CATCAAGTGTTCACATGCGGCAAGGAGGCGCACATCGCGCAACTGCGCGCCCGCCG 149
Db 46 CGTCAAGAAAGCGTACTTCAAGCGCTTCCAGACCAAGTTCCGTGCGCGCCCGGAGG 105
Qy 150 CAAGAAGGGTGTGTGAAGTCCGGCAAGGAGGCGCACAGATGAAGGACTCCACCGAGCG 209
Db 106 CAAGCGGACTACCGCGCGCCANCTGGTATCCAGGACAAAGAACTACACACGCC 165
Qy 210 CCAGGCCAACTTTCGCGGAGGACCTGGCTTCCCGCAGGCGCAAGCGCGCGAGTTCC 269
Db 166 CAAGTACCGTTGCTC-----GTCCGCTTCAACAACAGGAGACATCTCTCCAGATCG 219
Qy 270 CAGGAGCAGAACCGCGCAACAGCCCAACAGCGCGAGCTGCAGGTGCGCGCGACAA 329
Db 220 CTAGCCCAAGATCGACGGGACCAATCTCTGCGCGCGCTACTTCGCACGAGTCAACCG 279
Qy 330 CCCCAGCGAGCGCGCGCGCGAGCGCGGACACCTTGAACCTTCCCGCAGATCAACCT 389
Db 280 CTTGCGGTCAAGTCTGCGCTGACCAACTACGCGCGCGCTACGCGACTGCGCTGCT 339
Qy 390 GTGCAGCGCCCTGTGTAGCATCAAGGTGGGGGCGGACATCAAGAGGCGCTTGTGGA 449
Db 340 GCGCGCGCTGTCTGAAGAAGCTCAACCTTACTTCAAGTACGAGGTGTCAAGAAGT 399
Qy 450 CACCGCGCGGACACACCTGTGTGAGGAGATGAGCTTCCCGGAAGTGAAGGCCAA 509
Db 400 CAAGGGGAGACTACACCTCGAGGAGCTCGAGGCGCGCGCGCTTCAAGGCC-- 457
Qy 510 GATGATGCGCGCATCGCGGCTTCAATCAAGGTGCGCGAGTACGACAGATCTGTGATCGA 569
Db 458 -TGCTCGACGTGGCGCTGGTCCGACCTGAGCTGGCGCGCGGTGTTCGCGCGCTCAA 515
Qy 570 GATCTCGGCAAGAAGCCATCGGCACCTGCTGATCGGCGCGCCACCCCGTGAACATCAT 629
Db 516 GGGCATGTGCGGAGGCGGCTCAACGTCCCCACAGGAGACCGGCTTCGTGCGCTTCAA 575
Qy 630 CGGCGCAACATGCTGACCCAGCTGGGCTGCACCTGACCTTCCCATCAGCCCATCGA 689
Db 576 CGCGGACAGAAGAGTCAACGCCGCGGTCTCCGCAAGTACATCTTCGGCGCCACGT 635
Qy 690 GACCGTCCCGTGAAGTGAAGCCCGGCGATGGAGCGGCGCCCAAGGTGAAGCAGTGGCCCT 749
Db 636 CGCGCGGTACATGAAGCTCTCAAGGAGCAGGAGCGCGCGCTTCGACCGCGGCTTC 695
Qy 750 GACCGAGGAGAAGATCAAGCCCTGACCGCATCTGCGAGGAGATGAGAGAAGAGGCAA 809
Db 696 GCGCTACGCCAAGAGGAGTGTACCGCGGACATGCTCGAGAAGATCTACACCGAGCCCA 755
Qy 810 GATCACAAGATCGGCGCGCGAGAACCCCTACAACACCCCGCTGTTCGCCATCAAGAAGAA 869
Db 756 -----CAAGCAGATCGCGCGCGCGGACCTTTCGTCGCCCAAGCGCGCTTCGAGCCGA 809
Qy 870 GGACAGCACAAGTGGCGCAAGCTGTGTGAGTTCGCGGAGCTGAACAGAGCCACCCAGGA 929
Db 810 GGGCGCAAGCCCAAGCACTGGGGCAAGCGCAGGCTGACGTACCAAGAGGCAAGACCG 869
Qy 930 CTTCTGGGAGTCCAGTGGGATCCCCACCCCGCGCTCAAGAGAGAGAGAGCGGT 989
Db 870 CGTGGCGCGAGAGAGTCCGCTGGGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 929
Qy 990 GACCGTGTGAGGTGGGCGACGCTTACTTACGCGTGGCGCTTGGACGAGG 1039

Db 930 CGGTGCGCTGTACACCGCCATCTCCCGTGGCGTGTGCGGTGTCNGCG 979

RESULT 5

BM321430 867 bp mRNA linear EST 03-JAN-2002
 LOCUS rockefeller.0.1153 Mastigamoeba balamuthi lambda ZAP II Library
 DEFINITION Mastigamoeba balamuthi cDNA similar to ribosomal protein S4, mRNA sequence.

ACCESSION BM321430
 VERSION BM321430.1 GI:18055836

KEYWORDS EST.

SOURCE Mastigamoeba balamuthi.

ORGANISM Mastigamoebidae; Mastigamoebidae; Mastigamoeba.

REFERENCE 1 (bases 1 to 867)

AUTHORS Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and Philippe,H.

TITLE The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

MEDLINE 21819461

COMMENT Contact: Muller Miklos

Laboratory of Biochemical Parasitology

The Rockefeller University

1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu

Insert Length: 867 Std Error: 0.00

POLYA=No.

FEATURES

source

1..867

/organism="Mastigamoeba balamuthi"

/strain="ATCC 30984"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/notes="syn: Phreatamoeba balamuthi"

BASE COUNT 185 a 298 c 254 g 124 t 6 others

ORIGIN

Query Match 3.2%; Score 79; DB 13; Length 867;

Best Local Similarity 45.7%; Pred. No. 0.0012;

Matches 314; Conservative 0; Mismatches 370; Indels 3; Gaps 1;

QY 58 CAGCGCAGCACTTCAAGGCCCCAAGCGCATCATCAAGTGCITCAACTCGGCAAGGAG 117

Db 131 CCGCACAGATGCGCGAGTGCCTGCCGTGCATCTCTGTCGCAACAGGTTGAAGTAC 190

QY 118 GGCACATCGCCGCAACTGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGAAG 177

Db 191 GCGTGACCCGCGTGAGGTCACTCTGATGTCAGCGCCTGATCAAGATCGACGGC 250

QY 178 GAGGGCCACAGATGAGGACTGACCGGCGCAGGCCCAACTTCTTCGCGAGGACCTG 237

Db 251 AAGTCCGCGACCGACACAGCCTTCCCGCGGGTTCATGGACGTCGTCTCGATCGACAAG 310

QY 238 GCCTTCCCCAGGCGAGCGCGGAGTTCGCCAGCGCAGACACCGGCCAACAGCCCC 297

Db 311 ACCGACGAGCACTTCCGCCCTGTCTACGACACCAAGGGCGCTTCCAGGCGCACCGCATC 370

QY 298 ACCAGCGCGAGCTGCAAGGT---GCGGCGGACAAACCCCGCAGCGAGCGCGCGCGAG 354

Db 371 AACTGACCGAGGCCAAGTTCAAGTCGGCAAGTTCGCGCGGTGAGCTCGGCAACAAG 430

QY 355 CGCCAGGCGACCTTGAACTTCCCCAGATACCCCTGTGGCAGCGCCCCCTGGTGAGCATC 414

Db 431 GGCATCCGTAACCTGTGTGACGACGCGCGCAGCATCCGCTACCCCAACCCCGACATC 490

QY 415 AAGTGGCGCGCGCATCAAGGAGCCCTGCTGGACACCGCGCGCGCGACACACCGTCTG 474

Db 491 AAGGTCAACGACACGGTCAAGATCGACCTGCTCGGGCAAGATCATCGACTTCGTGAGG 550

QY 475 GAGGAGATGAGCTGCCCGCAAGTGGGAAGCCCAAGATGATCGGCGGATCGCGGGCTTC 534

Db 551 TTCCAGATCGCAACCTCGTCTCATGATCACTGGCGGACCACTTGGCGGCTCGGCGTC 610

QY 535 ATCAAGTGGCCCATGACGACCATGATCTGATCGAGATCTCGGGAAGAAGCCATCGGC 594

Db 611 ATTGTGCGCGCGAAGACGAGGCGTCTGATCGAGATCATCCAGTCAAGGACCGCGTC 670

QY 595 ACCGTGTGATGCGGCCCAACCCCGTGAACATCATCGCCCGCAACATGCTGACCCAGTCG 654

Db 671 GGCACCATGTTCCGACGCGCTGACCAACGCTTTCGTGATCGGCAAGGGACCAAGTCC 730

QY 655 GGTGTCACCTTGAATTCCTCCCATCGAGCCCATCGAGACCGTGCCTGAAGTGAAGCCC 714

Db 731 CTCGTACGCTGCCCGCGGCAAGGCATCAAGAAGTCATCATCGAGGAGTTCAGCGCG 790

QY 715 GGCATGACGCGCCCAAGGTGAAGCAG 741

Db 791 CGCCACGCGCACAGGACCGAGCAGGAG 817

RESULT 6

BM321022

LOCUS

DEFINITION

rockefeller.0.1192 Mastigamoeba balamuthi lambda ZAP II Library

Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC

3.3.1.1), mRNA sequence.

ACCESSION BM321022

VERSION BM321022.1 GI:18055428

KEYWORDS EST.

SOURCE Mastigamoeba balamuthi.

ORGANISM Mastigamoebidae; Mastigamoebidae; Mastigamoeba.

REFERENCE 1 (bases 1 to 1550)

AUTHORS Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and Philippe,H.

TITLE The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

MEDLINE 21819461

COMMENT Contact: Muller Miklos

Laboratory of Biochemical Parasitology

The Rockefeller University

1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu

Insert Length: 1550 Std Error: 0.00

POLYA=No.

FEATURES

source

1..1550

/organism="Mastigamoeba balamuthi"

/strain="ATCC 30984"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/notes="syn: Phreatamoeba balamuthi"

BASE COUNT 282 a 528 c 495 g 226 t 19 others

ORIGIN

Query Match 3.1%; Score 76.6; DB 13; Length 1550;

Best Local Similarity 45.6%; Pred. No. 0.0032;

Matches 308; Conservative 0; Mismatches 364; Indels 3; Gaps 1;

QY 1755 CGAGTGAAGCAGTGAACGCGCGTGCAGAAAGATCGCCATGGAGAGCATCGTGTCTG 1814

Db 5 CGGCACGAGGACGAGGACCAACCGCGCGCTCCATCGCGAGGCGCGTCTCGGTCTT 64

QY 1815 GGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAGAGACCTGGGAGACCTGGTGAC 1874

Db 65 CGCTGGAAGGCGGAGAACCTCCAGGAGTACTGGGAGTGCACCTGGGAAGGCCCTGTGCTT 124

QY 1875 CGACTACTGCGAGCCACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCGCCCTGGT 1934

Db 125 CGGCCCCCTACGAGGCGCTTCAGATCATCGTCGACGCGCGGTGACGCGACTCTGATGAT 184

ORIGIN		Query Match 3.0%; Score 75; DB 13; Length 862; Best Local Similarity 46.2%; Pred. No. 0.0057; Matches 293; Conservative 0; Mismatches 327; Indels 1;
QY	1935 GAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCTCGGCGCGAGACCTTCTACGTGGA	1994
Db	185 CCACAGGGGTTGCGCGCGAGGACACCCCAAGCTGCTGGAGGACGACGAGGCGCTCGA	244
QY	1995 CGGCGCGCCCAACCGGAGACCAAGATTCGGCAAGCGCGGTACGTGACCGACCGGGCGG	2054
Db	245 GGAGGTGCGCTGCTCAACACGTTTCAAGCAGTTCGAAGCAGGAGCAGCGCGCTTCTG	304
QY	2055 GCAGAGATCGTGAAGCTGACCGAGACCAACACCAAGAGACCGAGCTGACGCCATCCA	2114
Db	305 GCACAGATCCTCCCGAGATCCCGGCTGTGACGGAGGAGACGACACTGGCCTGATGAG	364
QY	2115 GCTGGCCCTCAGGACAGCGGCGAGGTGAACATCGTGACCGACGACCGACGACGCT	2174
Db	365 GCTGTACCACTGCACCGGACGCGACGTGCTGTTCCCGCGCTCAACGTCAACGACTC	424
QY	2175 GGGCATCATCCAGCGCCAGCCCGACAGAGCGAGAGCGAGCTGGTGAACAGATCATCGA	2234
Db	425 TG---TCACCAAGAGGAAGTTTGACACATCTACGGGTGCGGCCACTCGCTCATCGACG	481
QY	2235 CGAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGTGCCCGCCACACAGGCGATCGG	2294
Db	482 CATCAAGCGGCGGACCGACGTGATCTGCGCGCAAGGTGCGCGCTGCTCGCGGGCTACG	541
QY	2295 CGGCAACGAGCATCGACAAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTTCTCTGGA	2354
Db	542 CGACGTGGCAAGGGCTGCGCCGAGTGCTGCGCGCGGCGGCTGCGCGCTCATCTGTGAC	601
QY	2355 CGGCATCGATGGCGGATCGTGTATACCAAGTACATGGAGCAGCTGTACGTGGGCGAGCG	2414
Db	602 GGAGATCGACCCCATCTGCGCGTGCAGGGCGTGATGGCGGGCTTCGAGGTCAACAGCT	661
QY	2415 CGGCCCTAGGATCGA	2429
Db	662 CGAGGCGGGGCTCGA	676
RESULT 7		
LOCUS		BM321023 862 bp mRNA linear EST 03-JAN-2002
DEFINITION		rockefeller.0.594 Mastigamoeba balamuthi lambda ZAP II Library Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC 3.3.1.1), mRNA sequence.
ACCESSION		BM321023
VERSION		BM321023.1 GI:18055429
KEYWORDS		EST
SOURCE		Mastigamoeba balamuthi.
ORGANISM		Mastigamoeba balamuthi
REFERENCE		1 (bases 1 to 862)
AUTHORS		Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba. Bapteste, E., Brinkmann, H., Lee, J.A., Moore, D.V., Sensen, C.W., Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and Philippe, H.
TITLE		The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE		21819461
COMMENT		Contact: Muller Miklos Laboratory of Biochemical Parasitology The Rockefeller University 1230 York Avenue, New York, NY 10021, USA Email: mmulle@rockvax.rockefeller.edu Insert Length: 862 Std Error: 0.00 POLYA-No. Location/Qualifiers
FEATURES		1..862 /organism="Mastigamoeba balamuthi" /strain="ATCC 30984" /db_xref="taxon:108607" /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library" /note="syn: Phreatamoeba balamuthi"
BASE COUNT		175 a 294 c 260 g 126 t 7 others
LOCUS		BM321393 853 bp mRNA linear EST 03-JAN-2002
DEFINITION		rockefeller.0.1222 Mastigamoeba balamuthi lambda ZAP II Library Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA sequence.
ACCESSION		BM321393
VERSION		BM321393.1 GI:18055799
KEYWORDS		EST
SOURCE		Mastigamoeba balamuthi.
ORGANISM		Mastigamoeba balamuthi
REFERENCE		1 (bases 1 to 853)
AUTHORS		Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba. Bapteste, E., Brinkmann, H., Lee, J.A., Moore, D.V., Sensen, C.W., Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and Philippe, H.
TITLE		The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE		21819461
COMMENT		Contact: Muller Miklos Laboratory of Biochemical Parasitology

The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockefeller.edu
Insert Length: 853 Std Error: 0.00
POLYA=No.

FEATURES source
Location/Qualifiers
1..853
/organism="Mastigamoeba balamuthi"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"
BASE COUNT 156 a 307 c 263 g 125 t 2 others
ORIGIN

Query Match 3.0%; Score 73.2; DB 13; Length 853;
Best Local Similarity 46.9%; Pred. No. 0.011;
Matches 299; Conservative 0; Mismatches 329; Indels 9; Gaps 2;
QY 388 CTGTGGAGCGCCCTGGTGAGCATCAAGGTGGCGGCCAGATCAAGGAGGCCCTGTGCTG 447
Db 125 CTGGCCCGCGCTGTCTGAAGAAGCTCAACCTCGACTCCAAGTACGAGGGTGTCAAGAAG 184
QY 448 GACACCGCGCGGACGACACCGCTGTCTGAGAGATGAGCCTGCCCGGCAAGTGGAGGCC 507
Db 185 GTCAACGGCGGAGGACTACAACGTGAGAGCTCGACGACGGCGCCGCCCTTCAAGGCC 244
QY 508 AAGATGATCGCGCGCATCGCGCGCTTCATCAAGGTGGCGGCTAGTACGACCATCTGTATC 567
Db 245 CTGCTCGAGCTCGCGCTGTCTGCGACCTCGACGGCGCGCGGTTCGCGCGCCCTCAAG 304
QY 568 GAGATCTCGCGGAGAGAGCGCATCGGCACCTGTGTATGCGGCCCGCCCGCGTGAACATC 627
Db 305 GGCATGTGCGAGCGCGCGGTCAACGTCCCG---CACACGAGACCGCGTTCGTGCGCTTC 361
QY 628 ATCGGCCCGACATGCTGACCCAGCTGGCTGCGACCTGACCTTCCCATCAGCCCCATC 687
Db 362 AACGGCGACAAGAGGAGCTCAACGCCGCGTCTCCGCAAGTATCTTCGGCGGCCAC 421
QY 688 GAGACCTGTCGCGGTGAAGCTGAAGCGCGCATGAGCGCGCCCGCAAGGTGAACAGTGGCC 747
Db 422 GTCGCGGCTACATGAAGCTCTCAAGGAGGACGACGCGCGCGCTTCGACCGCGCATTC 481
QY 748 CTGACCGAGGAGAGATCAAGCGCCTGACCGCCATCTCGAGGAGATGGAGAAGGAGGC 807
Db 482 TCGCGCTACGCAAGAGAGGTGTACCGCGCGACATGCTCGAAGAGATCTACACCGAGGCC 541
QY 808 AAGATCACCAAGATCGGCGCGCGAGACCCCTACACACCCCGGTGTCGCCATCAAGAAG 867
Db 542 CA-----CAAGCAGATCCGCGCGCGACCCGACCTTCGTCGCCAAGCGCGCTCGAAGGCC 595
QY 868 AAGCAGACGACCAAGTGGCGCAAGCTGGTGACTTCGCGAGCTGAACAGCGCACCCAG 927
Db 596 GAGGCGCGCAAGCCAGCACTGGGCGAAGCGGAGGTGACGTACCAAGCGCGCAAGAAC 655
QY 928 GACTTCTGGAGGTGAGCTGGGGCATCCCCAACCCCGCGCCTGAAGAAGAAGAGAGC 987
Db 656 CGGCTCGCCAGAGAAGGTCCCGCTGGGTATCCCGACGCGCCCGCAAGAGCCATTAATT 715
QY 988 GTGACCGTGTGGAGCTGGGCGAGCGCTACTTACAG 1024
Db 716 CCGGTCCTCTGTACACNGCCATCTCCGCTGCGGCG 752

RESULT 9
BM587428/c 566 bp mRNA linear EST 25-FEB-2002
LOCUS 1700687321202 A.Gam.ad.cdna.blood1 Anopheles gambiae cDNA clone
DEFINITION 19600449696110 5', mRNA sequence.
ACCESSION BM587428
VERSION BM587428.1 GI:18883289
KEYWORDS EST.
SOURCE African malaria mosquito.

ORGANISM

Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.

REFERENCE

1 (bases 1 to 566)
Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab
, R., Collins, F.H., Venter, J.C. and Hoffman, S.L.
Celera Anopheles gambiae EST project
Unpublished (2002)

TITLE

JOURNAL

COMMENT

Celera Genomics
Contact: Holt R.A.
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU01004AAX row: G column: 08
Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers
1..566
/organism="Anopheles gambiae"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449696110"
/clone_lib="A.Gam.ad.cdna.blood1"
/dev_stage="Adult"
/lab_host="DH10b"
/note="Vector: pSport1; site_1: SalI; site_2: NotI; whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cDNA inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"
BASE COUNT 60 a 202 c 165 g 139 t
ORIGIN

Query Match 3.0%; Score 72.8; DB 13; Length 566;
Best Local Similarity 48.3%; Pred. No. 0.013;
Matches 203; Conservative 0; Mismatches 217; Indels 0; Gaps 0;
QY 1950 GCTGGAGAGAGGCCCATCATCGGCGCGGAGACCTTCTACGTGGACGGCGGCCAACCG 2009
Db 482 GTTCGAGATCGACGCCAACGGCATCTCGAGGTGTGCGCCGAGGACAGGGCACGGGCA 423
QY 2010 CGAGACCAAGATCGCAAGGGCGGTAGTGACCGACCGGGCGCGGAGAGATCGTGAG 2069
Db 422 CCGGAAAGATGCTCATCACCACGACGACGACCGCTGACGCGGAGCATCGAGCG 363
QY 2070 CCTGACCCGAGACCAACAGAGAGCGAGCTGCGAGGCCATCCAGCTGGCCCTCGAGA 2129
Db 362 CATGATCAAGATCGGAGCGGTTCCGCGCGACGACGACGAAGAGCTGAAGGAGCGGTGA 303
QY 2130 CAGCGGACGAGGTGAACATCGTGACCGACGACGACGACGACGACGACGACGACG 2189
Db 302 GGCCCGCAACGAGCTCGAGAGTACGCGCTACAGCTGAAGAACCAGCTCAGCTCGAAGA 243
QY 2190 CCAGCCCGACAGAGCGAGAGCGAGCTGGTGAACGAGATCATCGAGAGCTGATCAAGAA 2249
Db 242 CAAGCTGGCGGAGAGCTGTCGCGACGACGACGACGACGACGACGACGACGACGACG 183
QY 2250 GGAGAAGGTGTACCTGAGCTGGGTGGCGCGCCACAGAGGCCATCGCGCGCAAGAGCAGAT 2309
Db 182 GAAGATCAAGTGGTGGAGGAGAGACGAGGACGAGGACGAGGAGGAGGAGGAGGAGG 123
QY 2310 CGACAAGCTGTGAGCAGGAGCATCCGCAAGTGTGTCTCTGGAGCGCATCGATGGCGG 2369
Db 122 GAAGGAGCTGGAGAGACATCGTGCAGGCCCATCATTTGCCAAGCTGTACGCGAGCAGTGGCGG 63

RESULT 10

BE601575

LOCUS

DEFINITION

HVSMEH0098K08f

640 bp mRNA linear EST 22-OCT-2001
HVSMEH0098K08f Hordeum vulgare 5-45 DAP spike EST library


```
BASE COUNT      108 a 177 c 139 g 76 t
ORIGIN

Query Match
Best Local Similarity 2.9%; Score 71.6; DB 13; Length 500;
Matches 209; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 789 GGAGATGGAGGAGGAGGCAAGATCACCAAGATCGCGCCCGGAGAACCCCTACAAACACCCC 848
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 32 GGAGAGGTGACGGTGAAGGTGTCGGCCCAAGATGATCTCGGTGACGGGGCCCGCGGCAC 91
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 849 CGTGTTCGCCATCAAGAAGAAGGAGCAGCACCAGTGGCGCAAGCTGGTGGACTTCCGCGA 908
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92 CTTGACCCGCAACTTCAAGCACCTCAACCTCGACTTCAGCTGCGAGGACGGGGCGGCAA 151
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 909 GCTGAACAGCGCACCCAGGACTTCTGGAGGTGCGAGCTGGGCAATCCCGACCCCGCGG 968
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 GCTCAAGGTGACCGCTGTTTCGGCACCCGCGCACCATGCGCGGCATCCCGACCGGCAT 211
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 969 CTTGAAGAAGAAGAGAGCGTGACCGTCTGGAGCTGGCGGCGACGCTTACTTCAGGTC 1028
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 212 CTCCACCTCAGAACCTCATACCGGGCTCACCAGGGCTTCGCTACAAAGATCGGCTT 271
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1029 CTTGACAGGACTTCCGCAAGTACACCGCTTCAACATCCCGAGCATCAACAACGAGAC 1088
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 272 CGTCTACCTCCTTCCCGCATCAAGCGCTCCATCACCSCCGCCCAACCGGGGCATCGAGAT 331
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1089 CCCCAGGATCCGCTACAGTACAGGTGCTGCCCGAGGGCTGGAGGCGACCCCGAGCAT 1148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 332 CCGCAACTTCTCGCGGAGAGAGGTGAGGAGGTGGACATGCTCGACGGCAACGACATCGAGT 391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1149 CTTCCAGAGCATGACCAAGATCCTGGAGCCCTTCGCGCCGCGCAACCCGAGATGCT 1208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 392 CTTGCGGTCCGAGAGGTCAGAGTATGATGCTCTCGACGGCAACGACATCGAGTCTGT 451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1209 GATCTACAGCCGCCCT 1226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 452 CTCCCGTCCGCGGCCCT 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
BM368580
LOCUS
DEFINITION
    Ebem08_SQ004_A06_R embryo, 40 DPA, no treatment, cv Optic, Ebem08
    Hordeum vulgare cDNA clone Ebem08_SQ004_A06 5', mRNA sequence.
ACCESSION
    BM368580
VERSION
    BM368580.2 GI:21937722
KEYWORDS
    EST.
SOURCE
    Hordeum vulgare.
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
    ; Triticeae; Hordeum.
REFERENCE
    1 (bases 1 to 538).
    Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
    Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
    Development of Barley Transcriptome Resources
    Unpublished (2001)
    On Jan 10, 2002 this sequence version replaced gi:18111970.
    Contact: Waugh R, Marshall DF
    Genome Dynamics/Computational Biology
    Scottish Crop Research Institute
    Invergowrie, Dundee, DD2 5DA, Scotland, UK
    Tel: 00 44 1382 562731
    Fax: 00 44 1382 562426
    Email: est@scri.sari.ac.uk
    All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse.
Location/Qualifiers
    1..538
    /organism="Hordeum vulgare"
    /cultivar="Optic"

FEATURES
    source
```

```
/db_xref="taxon:4513"
/clone="EBem08_SQ004_A06"
/clone_lib="embryo, 40 DPA, no treatment, cv Optic,
EBem08"
/tissue_type="embryo"
/dev_stage="40 DPA"
/lab_host="DH108"
/Note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from embryos dissected from developing grains (40
days post anthesis) in glasshouse grown barley plants.
Developed as part of the barley transcriptome resources of
BBSRC/SEERAD funded cereal IGF (Investigating Gene
Function) project."
BASE COUNT      107 a 205 c 149 g 77 t
ORIGIN

Query Match
Best Local Similarity 2.9%; Score 71.6; DB 13; Length 538;
Matches 209; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 789 GGAGATGGAGGAGGAGGCAAGATCACCAAGATCGCGCCCGGAGAACCCCTACAAACACCCC 848
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 93 GGAGAGGTGACGGTGAAGGTGTCGGCCCAAGATGATCTCGGTGACGGGGCCCGCGGCAC 152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 849 CGTGTTCGCCATCAAGAAGAAGGAGCAGCACCAGTGGCGCAAGCTGGTGGACTTCCGCGA 908
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 153 CTTGACCCGCAACTTCAAGCACCTCAACCTCGACTTCCAGCTGCGAGGACGGGGCGGCAA 212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 909 GCTGAACAGCGCACCCAGGACTTCTGGAGGTGCGAGCTGGGCAATCCCGACCCCGCGG 968
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 213 GCTCAAGGTGACCGCTGTTTCGGCACCCGCGCACCATGCGCGGCATCCCGACCGGCAT 272
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 969 CTTGAAGAAGAAGAGAGCGTGCACCGTCTGGAGCTGGCGCGCAGCTTACTTCAGGTCGCC 1028
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 273 CTCCACCTCAGAACCTCATCACCAGGCTCACCAGGGCTTCCGCTACAAAGATCGGCTT 332
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1029 CTTGACAGGACTTCCGCAAGTACACCGCTTCAACATCCCGAGCATCAACAACGAGAC 1088
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 333 CGTCTACCTCCTTCCCGCATCAAGCGCTCCATCACCSCCGCCCAACCGGGGCATCGAGAT 392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1089 CCCCAGGATCCGCTACAGTACACGCTGCTGCCCGAGGGCTGGAGGCGACCCCGAGCAT 1148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 393 CCGCAACTTCTCGCGGAGAGAGGTGAGGAGGTGGACATGCTCGACGGGTCCACCAT 452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1149 CTTCCAGAGCATGACCAAGATCCTTGAGGCCCTTCGCGCCCGCCCAACCCCGAGATGCT 1208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 453 CTTGCGGTCCGAGAGGTCAAGGATGAGATGCTCTCGACGGCAACGACATCGAGTCTGT 512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1209 GATCTACAGCCGCCCT 1226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 513 CTCCCGTCCGCGGCCCT 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
AJ471121
LOCUS
DEFINITION
    AJ471121 S00008 Hordeum vulgare cDNA clone S0000800191A10F1, mRNA
    sequence.
ACCESSION
    AJ471121
VERSION
    AJ471121.1 GI:21187077
KEYWORDS
    EST.
SOURCE
    Hordeum vulgare.
    Hordeum vulgare.
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
    ; Triticeae; Hordeum.
REFERENCE
    1 (bases 1 to 540).
    Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
    Barley EST's
    Unpublished (2002)
    Contact: Schulman AH
    Institute of Biotechnology
```


SOURCE
ORGANISM

Hordeum vulgare.
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.

REFERENCE
AUTHORS

Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.

TITLE
JOURNAL

Development of Barley Transcriptome Resources
Unpublished (2001)

COMMENT

On Jan 10, 2002 this sequence version replaced gi:18120502.

Contact: Waugh R, Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk

All sequence has a Phred quality score of 20 or over

Seq primer: M13 reverse.

FEATURES
source

1. .578
Location/Qualifiers
/organism="Hordeum vulgare"
/cultivar="Optic"
/db_xref="taxon:4513"
/clone="EBem05_SQ004_C16"
/clone_lib="embryo, 14 DPA, no treatment, cv Optic,
EBem05"
/tissue_type="embryo"
/dev_stage="14 DPA"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site.1: Sal I; Site.2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from embryos dissected from developing grains (14
days post anthesis) in glasshouse grown barley plants.
Developed as part of the barley transcriptome resources of
BBSRC/SEERAD funded cereal IGF (Investigating Gene
Function) project."
BASE COUNT 116 a 224 c 155 g 83 t
ORIGIN

Query Match 2.9%; Score 71.6; DB 13; Length 578;

Best Local Similarity 47.7%; Pred. No. 0.02;

Matches 209; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 789 GGAGATGGAGAGAGGGCGGACAGATCACCAGATCGGGCCCGGAGAACCCCTACACACCCC 848

Db 114 GGAGGAGGTGACGTGAAGGTGTGCGGCAAGATGATCTCGGTGACGGGGCGCGCGGCAC 173

QY 849 CGTGTTCGCCATCAAGAGAGGACACCAAGTGGCGCAAGCTGGTGGACTTCCGCGA 908

Db 174 CCTGACCGCAACTTCAAGCACCTCACTCGACTTCCAGCTGAGGACGGGGGGCGAA 233

QY 909 GCTGAACAAGCGCACCGAGACTTCTGGGAGGTCCAGCTGGGCATCCCCACCCCGCGG 968

Db 234 GCTCAAGGTGACGCTGTGGTGGCGCGGACCACTGGCGGCCATCCGACCGCCAT 293

QY 969 CCTGAAGAAGAGAGAGCGTGACCGTGTGGAGCTGGGGAGCGCCCTACTTCAGCGTGCC 1028

Db 294 CTCACCGTCCAGAACCTCATACCGCGCGTCACCAAGGGGTTCGCGCTACAAGATCGCGTT 353

QY 1029 CCTGGACGAGACTTCGCGCAAGTACCGGCTTCACCATCCCGAGCATCAACACAGAGAC 1088

Db 354 CGTCTAGCTCACTTCCCGCATCAGCGCTTCATCACCAGCGCCCAACCGGGGCATCGAGAT 413

QY 1089 CCCCGGCATCCGCTACAGTACACAGTCTGCTGCCAGGGGTGGAAGGCGAGCCCGCAGCAT 1148

Db 414 CCGCAACTTCTCGCGGAGAGAGGTGAGGAAGTGGACATGCTCGACGGGGTCAACCAT 473

QY 1149 CTTCCAGAGAGCATGACCAAGATCTCTGGAGCCCTTCGGGCGCGCGCAACCCCGAGATCGT 1208

Db 474 CTTGCGGTCCGAGAGGTCAAGATGATGATGCTCTCGAGCGCAACGACATCGAGCTCGT 533

QY 1209 GATCTACCAGGGCCCCCT 1226

Db 534 CTCCCGCTCGCGCGCCCT 551

Search completed: February 10, 2003, 20:43:40

Job time : 3399 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:14:58 ; Search time 80 Seconds
(without alignments)
9441.811 Million cell updates/sec

Title: US-09-610-313-31
Perfect score: 2463
Sequence: 1 gtcgacgaccatgcccga.....gggtagcaccggtgaattc 2463

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 1533381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1106.8	44.9	2601	4	US-09-117-217-7
2	1106.8	44.9	2601	4	US-09-117-217-9
3	1106.8	44.9	2601	4	US-09-117-217-11
4	1106.8	44.9	2601	4	US-09-117-217-13
5	1100.4	44.7	7399	2	US-08-418-848A-9
6	1100.4	44.7	9709	2	US-08-188-583-5
7	1100.4	44.7	9709	3	US-08-388-353-1
8	1100.4	44.7	9709	3	US-08-488-551B-1
9	1100.4	44.7	9709	4	US-09-309-572-15
10	1100.4	44.7	12494	4	US-08-935-312-13
11	1100.4	44.7	12494	4	US-08-848-760B-33
12	1100.4	44.7	15581	3	US-08-646-538-35
13	1100.4	44.7	15581	4	US-09-503-222-35
14	1090.8	44.3	9737	2	US-08-944-449-7
15	1090.8	44.3	9737	4	US-09-353-362-7
16	1055.6	42.9	9746	1	US-08-022-835-3
17	1055.6	42.9	9746	1	US-08-388-809-3
18	1055.6	42.9	9746	2	US-08-647-714-3
19	1054.4	42.8	8932	4	US-09-124-900-1
20	1054.4	42.8	8933	3	US-08-463-210-4
21	1054.4	42.8	8933	4	US-09-620-958A-3
22	1054.4	42.8	8933	4	US-09-620-958A-4
23	1054.4	42.8	8933	4	US-09-620-958A-9
24	1042	42.3	9739	1	US-08-022-835-1
25	1042	42.3	9739	1	US-08-388-809-1
26	1042	42.3	9739	2	US-08-647-714-1
27	1036.8	42.1	5362	3	US-08-463-210-5

28	1025	41.6	3033	2	US-07-743-357-19	Sequence 19, Appl
29	1025	41.6	4113	2	US-07-743-357-21	Sequence 21, Appl
30	1019.2	41.4	9207	3	US-08-388-353-800	Sequence 800, App
31	1019.2	41.4	9207	3	US-08-488-551B-800	Sequence 800, App
32	929.4	37.7	3856	2	US-07-743-357-20	Sequence 20, Appl
33	924	37.5	2739	2	US-07-743-357-16	Sequence 16, Appl
34	843.6	34.3	2348	2	US-08-876-546A-1	Sequence 1, Appl
35	843.6	34.3	2348	4	US-09-412-252-1	Sequence 1, Appl
36	843.6	34.3	2348	4	US-09-079-675-1	Sequence 1, Appl
37	831	33.7	9793	1	US-08-470-202-56	Sequence 56, Appl
38	831	33.7	9793	1	US-08-471-770-56	Sequence 56, Appl
39	831	33.7	9793	2	US-08-468-059-56	Sequence 56, Appl
40	831	33.7	9793	4	US-09-109-916-56	Sequence 56, Appl
41	823.4	33.4	1256	5	PCT-US93-06748-1	Sequence 1, Appl
42	604	24.5	1256	4	US-09-158-695-18	Sequence 18, Appl
43	553	22.5	3168	2	US-08-659-251-14	Sequence 14, Appl
44	553	22.5	3168	4	US-09-256-490-14	Sequence 14, Appl
45	553	22.5	3168	5	PCT-US96-11445-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-117-217-7
; Sequence 7, Application US/09117217
; Patent No. 6221578
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117.217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(492)
; OTHER INFORMATION: gag Polyprotein
US-09-117-217-7

Query Match	44.9%	Score 1106.8;	DB 4;	Length 2601;
Best Local Similarity	68.2%	Pred. No. 1.4e-173;		
Matches 1618;	Conservative	0;	Mismatches 732;	Indels 22; Gaps 5;
QY	14	TGGCCGAGGCGCATGAGCGAGCCACCA---GCGCCAAATCCTGATGAGCGCAGCAACT	70	
Db	77	TGGCTGAAGCAATGAGCCAAATTAACAAATTCAGCTACCAATATGATGAGAGAGCAATT	136	
QY	71	TCRAAGGCCCCAAGCGCATCATCAAGTCTCACTGCGGCAAGGAGGCGCCATCGCCC	130	
Db	137	TTAGAACCAAGAAAGATTGTTAAGTGTTCATTTGGCAAGAGAGGCGCACACAGCCA	196	
QY	131	GCAACTTCGCGCGCCCCCGCAAGAGGCTGCTGGAAGTGGCGCAAGGCGCCACACAGA	190	
Db	197	GAATTGCGAGGCCCCCTAGGAAAGAGGCTGTTGGAATGTGGAAGAGGAGACACAAA	256	
QY	191	TGAAGACTGCACCGAGCGCCAGCCAACTTCTTCGCGAGGAGCTGCCCTTCCCCCAGG	250	
Db	257	TGAAGATTGTTACTGAGACAGACAGCTAA-TTTTTTAGGGAAGATCTGCCCTTCTTACAAG	315	
QY	251	GCAAGGCCCCGAGTTCCTCCAGCGAGACAGACCGCCCAACAGCCCCCAGCCCGAGC	310	
Db	316	GAAAGGCGCGGAATTTTCTTCAGAGCAGACAGCCAGCCAGCCACAGAGAGAGAGC	375	

1876	Ddb	CCACAGAAGCATAGTAATATGGGAAGAGACTCCTAAATTTAAACTGCCCATACAAGG	1935
1853	Qy	AGACCTGGGAGACACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGT	1912
1936	Ddb	AAACATGGGAACATGGTGGACAGATATTGGCAAGCCACCTGGATTCCTGACTGGGAGT	1995
1913	Qy	TCGTGAACACCCCCCCCCTGGTGAAGCTGTGGTACCAAGCTGGAGAGAGGCCCATCATCG	1972
1996	Ddb	TTGTTTAATACCCCTCCCTTAGTGAATATATGGTACCAAGTTAGAGAAAGAACCCATAGTAG	2055
1973	Qy	GGCGGAGACCTTCTACGTGACGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCCG	2032
2056	Ddb	GAGCAGAACCCTTCTATGTAGATGGCGACGTAAACAGGAGACTAAATTTAGGAAAAGCAG	2115
2033	Qy	GCTACCTGACCGACCGGGCGCGCAGAGAATCGTAGCCCTGACCGAGACCACCAACACG	2092
2116	Ddb	GATATGTTACTAATAGAGGAAGACAAAAGTTGTCACCCTAACTGACACAAACAATCAGA	2175
2093	Qy	AGACCGAGCTGGAGCCATCAGCTGGCCCTGCAGGACACGGCGAGCGAGGTGAACATCG	2152
2176	Ddb	AGACTGAGTTACAAGCAATTTATCTAGCTTTGCGAGATTCCGGGAATTAGAAGTAAACATAG	2235
2153	Qy	TCACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCACCCCGACAAAGAGCGAGAGCG	2212
2236	Ddb	TAAACAGCTACAATATGCATTAGGAATCATTAAGCACACACAGATCAAGTGAATCAG	2295
2213	Qy	AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTTGTAACCTGAGCTGGG	2272
2296	Db	AGTTAGTCAATCAAAATAATAGACGAGTTTAATAAAAAAGGAAAAGGTCTATCTGGCATGGG	2355
2273	Qy	TGCGCGCCACAAAGGCGATCGCGCGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGCA	2332
2356	Db	TACCAGCACAAAGGAATTTGGAGGAATGAACAAAGTAGATAAAATTTAGTCAGTCTGCGAA	2415
2333	Qy	TCCGCAAGGTGCTGTCTCTGGACGGCATCGAT	2364
2416	Db	TCAGGAAGTACTATTTTTAGATGGAATAGAT	2447

RESULT 3

```

US-09-117-217-11
;
; Sequence 11, Application US/09117217
; Patent No. 6221578
;
; GENERAL INFORMATION:
;
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Rudi
; APPLICANT: FAUWELS, Rudi
;
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
;
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
;
; FILE REFERENCE: 1377-125P
;
; CURRENT APPLICATION NUMBER: US/09/117,217
;
; CURRENT FILING DATE: 1998-07-24
;
; NUMBER OF SEQ ID NOS: 15
;
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 11
;
; LENGTH: 2601
;
; TYPE: DNA
;
; ORGANISM: HIV-HXB2
;
; FEATURE:
;
; NAME/KEY: CDS
;
; LOCATION: (453)..(749)
;
; OTHER INFORMATION: Protease
;
US-09-117-217-11

```

Qy	71	TC	AAGGCGCCCAAGCGGATCATCAAGTCTCTCAACTGCGGCAAGGAGGGCCACATCGCCC	130
Db	137	TT	AGAACCAAGAAGAAAGATTGTTAAGTGTTCAAATGTGGCAAGAAGGGCAGACAGCA	196
Qy	131	GC	AACTCGCGCGCCCGCCGCAAGAAGGCGTCTGGAAGTGGCGAAGGAGGCCACACAGA	190
Db	197	GA	ATTGCGAGGGCCCTTAGGAAAAAGGCGTGTGGAAATGTGGAAAGGAGACACACAA	256
Qy	191	TG	AAGGACTGCACCGAGCGCCAGCCCAACTTCTTCCGCGAGGAAGTGGCTTTCCCCAGG	250
Db	257	TG	AAAGATTGTACTCAGAGACAGGCTAA-TTTTATTAGGAAAGATCTGCGCTTCTCTACAG	315
Qy	251	GC	AAGGCGCGGAGTTCCTCCACGAGCAGAGAACCGCGCCACAGCCCCACACGCGCGAGC	310
Db	316	GG	AAAGGCGAGGGAATTTTCTTCAGAGCAGACAGAGCCCAACAGCCCCACCAAGAGAGC	375
Qy	311	TG	CAGTGTGCGGG-----CGACAACCCCGCAGCGAGCGCGCGCCGAGCGCCAGGCA	364
Db	376	TT	CAGGCTGGGTAGAGACAACAACCTCCCTTCAGAAGCAGGAGCCGATAGACAAGAA	435
Qy	365	-----	CCCTGAATTCCTCCCGAGATCACCTGTGTGGCAGCGGCCCTGTGTGACCATCAAGG	418
Db	436	CT	GTATCTTTAACTTCCTCAGGTCACTCTTTGGCAACAGCCCCCTCGTCACAATAAAGA	495
Qy	419	TG	GCGGCACAGATCAAGGAGGCGCTGTGTGGACACCGCGCCGACACACCGTGTCTGGAGG	478
Db	496	TAG	GGGGCACTTAAGGAGAGCTATTAGATACAGGAGCAGATGATACAGTATTAGAG	555
Qy	479	AG	ATGAGCTTCGCGGGCAAGTGGAAAGCCCAAGATGATCGGGGCAATCGCGCGCTTCATCA	538
Db	556	AA	ATGAGTTTCCAGGAAGATGGAACCCAAAAATGATAGGGGAAATGGAGGTTTTATCA	615
Qy	539	AG	TGCGCCAGTACGACACAGATCTGTATCGAGATCTGGGGCAAGAGGCCATCGCACCG	598
Db	616	AA	GTAAAGACAGTATGATCAGATACTCATAGAAATCTGTGGACATTAAGGCTATAGGTACAG	675
Qy	599	TG	CTGATCGGGCCCAACCCCGCTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGT	658
Db	676	TA	TATAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTT	735
Qy	659	GC	ACCTGAACTTCCCATCAGCCCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCGCA	718
Db	736	GC	ACTTTAAATTTTCCCATTAGCCCTATTAGAGACTGTACCACTAAATTTAAAGCCAGGAA	795
Qy	719	TG	CAGCGGCCCAAGGTGAAGCAGTGGCCCTCGACCGAGGAGAAGATCAAGGGCCCTGACCG	778
Db	796	TG	ATGGCCCAAAAGTTAAACATGGCCATTGACAGAGAAATAATAAGCATTAGTAG	855
Qy	779	CA	TCTCGGAGGAGATGGAGAAGGAGGCGAGATCACCAGATTCGCGCCCGAGAACCCCT	838
Db	856	AA	ATTGTACAGAGATGGAAGAAGGAGGAAATTTCAAAAAATTTGGCGCTGAAAAATCCAT	915
Qy	839	ACA	CACCCCGGTGTTCGCCATCAAGAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGG	898
Db	916	AA	ATACTTCCAGTATTGGCCATAAAGAAAAAGACAGTACTTAATGGAGAAATAATAGTAG	975
Qy	899	ACT	TCCGGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGCGAGCTGGGCATCCCC	958
Db	976	AT	TTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATAACCA	1035
Qy	959	AC	CCCGCGGCTGGAAGAAGAAGAGCGTGACCGTGTGACGTGGCGGAGCGCCTACT	1018
Db	1036	AT	CCCGCAGGCTTAAAGAAAGAAAAATCAAGTAAACAGTACTGGATGTGGGTGATGCATATT	1095
Qy	1019	TC	AGGTGCCCTTGACGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCA	1078
Db	1096	TT	TAGTCCCTTAGATGAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAA	1155
Qy	1079	ACA	AGAGACCCCGGCATCCGCTACCAAGTGTGCCCCAGGGCTGGAAGGCCA	1138
Db	1156	ACA	ATGAGACACCGGATTAGATATCAGTACAAATGTGCTCCACAGGATGGAAGGAT	1215

Qy	479	AGATGACGCTTCCCGGCAAGTGGAAAGCCCAAGATGATCGGCGGCATTCGGCGGCTTTCAACA	538
Db	556	AAATGAGTTTTGCCAGAAGATGGAAACCACAAAATGATAGGGGAATTGGAGGTTTTATCA	615
Qy	539	AGGTGCGCCAGTAGCAGCACAGATCCTGATCGAGTCTGGCGCAAGAGCCCATCGGCACCG	598
Db	616	AAGTAAGACAGTATGATCGATACTCATAGAATCTGTGGACATAAAGCTATAGGTACAG	675
Qy	599	TGCTGATCGGCCCCACCCCGTGAAACATCATCGGCGCAACATGCTGACCCAGCTGGGCT	658
Db	676	TATTAGTAGGACCTACACCTGTCAACATAATTGGAGAAATCTGTTGACTCAGATTGGTT	735
Qy	659	GCACCCTGAATTTCCGCATACGCCCATCGAGACCGTGCCTGTGAAGTGAAGCCCAGCA	718
Db	736	GCACCTTTAAATTTTCCCATTTAGCCCTATTGTGACTGTACCACTAAAAATAAAGCCAGAA	795
Qy	719	TGGAGCGGCCCAAGGTGAAGCAGTGGCCCTCACCGAGGAGAAGATCAAGCGCCTGACCG	778
Db	796	TGGATGCCCCAAAGTTAAACATNGCCATTGACAGAGAAAAATAAAGACATTAGTAG	855
Qy	779	CCATCTCGGAGGAGATGGAGAAGGAGGGCAAGTACCAGAATCGGCCCGCCAGAACCCCT	838
Db	856	AAATTTCTACAGAGATGGAAAAGGAGCGAAAATTTCAAATAATTGGGCTTGAAATCCAT	915
Qy	839	ACAACACCCCGTTTCGCCATCAAGAAGACGACAGCACCAAGTGGCGCAAGCTGGTGG	898
Db	916	ACAATCTCCAGTATTGGCCATTAAGAAAAAAGACAGTACTAAATGGAGAAAAATTAGTAG	975
Qy	899	ACTTCCGCGAGCTGAACAAGCGCACCGAGCTTCTGGAGGTGCAGCTGGSCATCCCCC	958
Db	976	ATTTACAGAGACTATFAGAGACTCAGACTTCTGGGAAGTTCAATTAGGAATACCAAC	1035
Qy	959	ACCCCGCGGCTTGAAGAAGAAAGAGCGTGACCGTCTGGACOTGGCGACCGCTACT	1018
Db	1036	ATCCCGAGGCTTAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATT	1095
Qy	1019	TCAGCGTCCCTTGGAGGAGACTTCCGCAAGTACACCGCTTACACATCCCCAGATCA	1078
Db	1096	TTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATPACTGCAATTTACCATCTAGTATAA	1155
Qy	1079	ACAACGAGACCCCGGCATCCGCTACCACTCAACGCTGCTGCCCGAGGCTTGAAGGGCA	1138
Db	1156	ACAATGAGACACCAGGATTTAGTATCAGTACATGTCTTCCACAGGATGGAAAGAT	1215
Qy	1139	GCCCCAGCATTTCCAGAGCAGATGACCAAGATCCTGGAGCCCTTCCGGCCCGCAACC	1198
Db	1216	CACGAGCAATATTCCAAGTAGCATGACAAAAATCTTAGAGCCTTTAGAAAAACAAATC	1275
Qy	1199	CCGAGATCGTATCTACCA- ----GGCCCCCTGTAGCTGGGACGAGCCTGGAGATCG	1252
Db	1276	CAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATPCTGACTTGAATAAG	1335
Qy	1253	GCCAGCACCGCGCAAGATCGAGGAGCTCGCAAGACCTCTCGCTGGGTGGGGCTTACCA	1312
Db	1336	GGCAGCATAGAACAAAAATAGAGAGCTGAGACACATCTCTTGAGGTGGGACTTACCA	1395
Qy	1313	CCCCGACAAGAGCAACAGAGAGACCCCTTCTCTGTGATGGGTACAGCTGCAC	1372
Db	1396	CACGACACAAAAACATCAGAAGAACCCTCCATTCTTTGGATGGGTATGAACTCCATC	1455
Qy	1373	CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAAGGAGCTGGACCGTGAACG	1432
Db	1456	CTGATTAATGGACAGTACAGCCTTAGTGTCTGCCAAAAAGACAGCTGCAGCTGTCAATG	1515
Qy	1433	ACATCCAGAAGCTGGTGGCAAGCTGAACTGGGCGAGCCGATCTACCCGGGATCAAGG	1492
Db	1516	ACATACAGAAGTTAGTGGGAAATTTGAATTGGCAAGTCAGATTTACCGAGGATTAAG	1575
Qy	1493	TGCGCCAGCTGTGCAGCTGTCTGCGCGGCGCAAGGCCCTGACCGCATCTGTGCCCTGTA	1552
Db	1576	TAGGCAATTTGTAACTCTTTAGAGAAGCAAGCAAGCACTAACAGAAGTAAATACCCTAA	1635

QY	1553	CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTTGCGCGAGCCCGTGCACG	1612
Db	1636	CAGAAGAAGCAGAGCTAGAACTGGCAGAAACAGAGAGATTCTTAAAGAACCAACGATACATG	1695
QY	1613	GGTGTTACTACGACCCACGACGACCTGGTGCCCGAGATCCAGAACGACGGGCCACGACC	1672
Db	1696	GAGTGTATTATGACCATCAAAAGACTTAATAGCAAAATACAGACGAGGGGCAAGGCC	1755
QY	1673	AGTGGACCTTACCAGATCTACCAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA	1732
Db	1756	AATGGACATATCAAAATTATCAAGACCCATTTAAAAATCTGAAGAACGNAATATGCAA	1815
QY	1733	AGATGGCAGCCGCCACACCAAGAGCTGAAGCAGCTGACCGAGGCGGTGCAGAGATCG	1792
Db	1816	GAATGAGGGTGCCCACTAATGATGTAAAACAATTAACAGAGGCGAGTGCAAAAAATAA	1875
QY	1793	CCATGGAGAGCATCGTGATCTGGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAAGG	1852
Db	1876	CCACAGAAAGCATAGTAATATGGGGAAGAGCTCTAAATTTAAACTTGCCCTTACAAGG	1935
QY	1853	AGACCTGGGAGACCTGGTGGACCGACTACTTGGCAGGCCACCTGGATCCCCGAGTGGGAGT	1912
Db	1936	AAACATGGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTTGGATTCTCTGAGTGGGAGT	1995
QY	1913	TCGTGAACACCCCCCTGGTGAAGCTGTGTACCAGCTGGAGAGAGGACCCCATCATCG	1972
Db	1996	TTGTGTAATACCCCTCCCTTAGTGAATATATGTGTACAGTTGAGAGAAAGAACCCCATAGTAG	2055
QY	1973	GGCGCGAGACCTTCTACGTGGACGGCGGCCCAACCGCAGACCAAGATCGGCAAGGCCG	2032
Db	2056	GAGCAGAAACCTTCTATGTAGATGGGCGAGCTTAACAGGGAGACTAAATTAGGNAAGCAG	2115
QY	2033	GCTACGTGACGACCGGGCGCGCAGAGATCGTGAGCCTTGACCGAGACACCAACACAGA	2092
Db	2116	GATATGTTACTAATAGAGGAAGACAAAAAGTTGTCAACCCTAACTGACACAAACAAATCAGA	2175
QY	2093	AGACCGAGCTGCAGGCCATCCAGCTTGGCCCTCGAGACAGCGGCACGAGGTGAACATCG	2152
Db	2176	AGACTGAGTTACAGCAATTTATCTAGCTTTTCAGAGTTTCGGGATTTAGAAGTTAAACATAG	2235
QY	2153	TGACCCAGACCGAGTACGCCCTGGGCATCATCCAGGCCCGCCGCAAGAGCGAGAGCG	2212
Db	2236	TAAACAGCTCACAATATGCATTAGGAATCATTCAAGCAACACAGATCAAGGTGAATCAG	2295
QY	2213	AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGNAGAGAGAGGTGTACCTGAGCTGGG	2272
Db	2296	AGTTAGTCAATCAATAATAGAGCAGTTAATAAAAAGGAAAGGCTTATCTGTGGCATGGG	2355
QY	2273	TGCCCGCCCAAGGGCATCGCGGCGACGAGCATCGCAAGCTGGTGAGCAAGGGCA	2332
Db	2356	TACCAGCACAAAGGAATTGGAGAAATGAACAAGTAGATAAATTAGTCAGTGCTGGAA	2415
QY	2333	TCCGCAAGTGCTGTTCTCTGGACGCATCGAT	2364
Db	2416	TCAGGAAAGTACTATTTTATAGATGGAATAGAT	2447

RESULT 5
US-08-418-848A-9
; Sequence 9, Application US/08418848A
; Patent No. 5847096
; GENERAL INFORMATION:
; APPLICANT: SCHUBERT, MANFRED, HARMISON II,
; APPLICANT: GEORGE G., CHANG-JIE, CHEN, BANJERJEA, AKHIL
; TITLE OF INVENTION: DEFECTIVE, INTERFERING
; TITLE OF INVENTION: HIV PARTICLES
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L. L. P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10154

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/418,848A
; FILING DATE: 07-APR-1995
; CLASSIFICATION: 526
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/936,849
; FILING DATE: 28-AUG-1992
; CLASSIFICATION: 526
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-40910S2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7399 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-418-848A-9

```

```

Query Match      44.7%; Score 1100.4; DB 2; Length 7399;
Best Local Similarity 68.0%; Pred. No. 1.7e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;

```

```

QY 14 TGGCCGAGGCGCATGAGCCAGGCGACCA--CGGCCACATCTCTGATCGAGGCGAGCAACT 70
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1838 TGGCTGAAGCAATGAGCAAGTACCAATCCAGTACGATAGTACATGATACAGAAAGGCAATT 1837

QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTCTCAACTCGCGCAAGGAGGCGCACATCGCCC 130
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1898 TTAGGAACCAAGAAAGACGTTTAAGTGTTCATTTGTGCAAGAGAGGCGACATGCCA 1957

QY 131 GCAACTGCCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGCAGCAAGAGGGCGCCACAGA 190
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1958 AAAATTGCGAGGCGCCCTAGGAAAAGAGGCTGTGGAATGTGGAAGAAAGGACACCAAA 2017

QY 191 TGAAGGACTGCACGCGCGCCAGGCGCAACTTCTCCGCGAGGACCTGGCCCTTCCCCCAGG 250
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2018 TGAAGATTTGACTGAGAGACAGGCTAA-TTTTGTAGGGAAGATCTGGCCCTTCCCAAG 2076

QY 251 GCAAGGCGCGAGTTTCCCGAGGAGCAGAAACCGCGCAACAGCCCGCCAGCGCGGAGC 310
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2077 GGAAGGCGAGGAATTTCTTCAGAGCAGACAGGCGCCAGCGCCCGCCAGAGAGAGC 2136

QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGCGAGCGCGCCAGGCA 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2137 TTCAGGTTTGGGGAAGAGAGAACAACTCCCTCTCAGAGAGGAGGAGCGATAGAGAGAA 2196

QY 365 -----CCCTGAATCTCCCGAGATCACTCTGTGGCAGCGCCCTCGTGAGCATCAAGG 418
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2197 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAGA 2256

QY 419 TGGCGCGCCAGATCAAGAGGCGCTCTGACACCGCGCGCGAGCAGACCGCTGCTGGAGG 478
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2257 TAGGGGCGCAATTAAGAGAGCTCTATTAGATACAGAGCAGATGATACATGATTAGAG 2316

QY 479 AGATGAGCCTCGCCGCGCAAGTGAAGCCCAAGATGATCGGCGGATCGGCGGCTTCATCA 538
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2317 AAATGAATTTGCCAGGAGATGGAACCAAAATGATAGGGGGAATGAGGTTTATCA 2376

QY 539 AGTGCGCCAGTAGGACGAGATCTGATCGAGATCTCGGCAAGAGGCCATCGGACCG 598
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2377 AAGTAGGACAGTAGATCAGATCTATAGAAATCTCGGACATCAAGCTATAGGTACAG 2436

```

```

QY 599 TGCTGATCGGCCCGCCCGTGAACATCATCGGCCCAACATCTGACCCAGCTGGCT 658
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2437 TATTAGTAGGACCTACACCTGTCAACATATTGGAAGAAATCTGTCACACATTTGGCT 2496

QY 659 GCACCCCTGAATCTCCCATCAGCCCCATCGAGACCGTGCOCGTGAAGCTGAAGCCCGCA 718
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2497 GCACCTTTAAATTTTCCCATAGTCTCTATTGAGACTGTACCAAGTAAATTAAGCCAGAA 2556

QY 719 TGGAGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGGAGGAGAAATCAAGGCCCTGACC 778
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2557 TGGATGCCCCAAAGTTAAACAATGGCCATTGACAGAGAAATAAAGCATTAGTAG 2616

QY 779 CCATCTCGAGGAGATCGAAGAGGAGGCGCAAGATCAGCAAGATCGGCCCGCCAGAACCCCT 838
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2617 AATTTGTACAGAAATGGAAGAGGAAATTTCAAAATTTGGCCTGAAATCCAT 2676

QY 839 ACAACACCCCGTGTTCGCATCAAGAAAGAGACAGCAACCAAGTGGCGCAAGTGTGG 898
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2677 ACAATACTCCAGTATTTGCCATAAAGAAAGACAGTACTAAATGGAGAAATTAGTAG 2736

QY 899 ACTTCCCGGAGCTCAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGCATCCCC 958
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2737 ATTTTCAGAGAACTTAATAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCAC 2796

QY 959 ACCCGCGCGCTGAAGAAGAGAGCGGTGACCGTCTGAGCTGGCGGCGACGCTACT 1018
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2797 ATCTGCGAGGTTAAACAGAGAAATCAGTAACAGTACTGATGTGGCGCATGCATATT 2856

QY 1019 TCAGCTGCGCCTGGAGGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCA 1078
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2857 TTTTCAGTCTCCCTTAGATAAAGACTTCAGGAAGTATAGTGCATTTACCATACCTAGTATAA 2916

QY 1079 ACAACGAGACCCCGGCGATCCGCTACCAGTACAGCTGCTCCCGCAGGCTGGAGGCA 1138
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2917 ACAATGAGACACCGAGGATTAGATATCAGTACAATGTCTTCCACAGGATGGAAAGAT 2976

QY 1139 GCCCGACGATCTCCAGAGCAGCATGACCAAGATCTCGAGCGCTTCGCGCCCGCAACC 1198
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2977 CACCAGCAATATTCAGGTGATGATGACAAATCTTAGAGCCCTTTAGAAAACAAATC 3036

QY 1199 CCGAGATCGTATCTACCA-----GGCCCCCTGTAGCTGGGCGAGCACCTCGAGATCG 1252
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3037 CAGACATAGTCACTATCAATACATGATGATGATGATGATGATGATGATGATGATGATGAT 3096

QY 1253 GCGAGCAGCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGTGGGCTTCACCA 1312
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3097 GGCAGATAGAACAAATAGAGGAAGTACAGCAACATCTTGTGAGTGGGATTTACCA 3156

QY 1313 CCCCCGACAGGAGCAGGAGGAGGCGCCCTTCTGTGATGGCTTACGAGCTGCACC 1372
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3157 CACCAGACAAAACATCAGAAAGAACCTCCATTCCTTGGATGGGTTTGAATCCCATC 3216

QY 1373 CCGACAAGTGACCGTGCAGCCCATCGAGCTGCGCGAGAGAGAGAGCTGGACCGTGAACG 1432
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3217 CTGATAATGACAGTACAGCTATAGTGTCTGCCAGAAAGAGACAGCTGAGCTGTCAATG 3276

QY 1433 ACATCCAGAGCTGCTGGGCAAGTGAACCTGGCGCAGCAGATCTACCCCGCATCAAGG 1492
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3277 ACATACAGAAATAGTGGGAAATTAATTTGGCAAGTACAGATTTATCGAGGATTAAG 3336

QY 1493 TGGCCGAGCTGTGAAGCTGCTGGCGGCGCAAGGCGCTGACCGACATCTGTCGCCCTGA 1552
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3337 TAAGGCAATTTATGTAATCTTCTTGGGGAACCAAGGACTAACAGAACTAGTACCACATA 3396

QY 1553 CCGAGGAGCGCGAGCTGGAGCTGGCCGAGAACCGGAGATCTGCGCGAGCGCCGTGCACG 1612
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3397 CAGAAGAGCAGAGCTAGAACTGGCAGAAAACAGGGAGATTTCTAAAGAACCGGTACATG 3456

QY 1613 GCGTGTACTACGACCCAGCAAGGACCTGTTGGCGGAGATCCAGAAAGCGGCGCACAGC 1672
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3457 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAAAATACAGAAAGCGGGGCAAGGCC 3516

QY 1673 AGTGGACCTACAGATCTACCGAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1732

```


QY 719 TGGAGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCG 778
DB 2596 TGGTGGCCCCAAAAGTTAAACAAATGGCCATTGACAGAGAAAATAAAGCAATTAGTAG 2655
QY 779 CCATCTGGAGAGNATGAGAGAGGAGGACAGATCACCAGATCGGCCGCCGAGAACCCCT 838
DB 2656 AAATTTGTACAGAAATGAAAGGAAGGAGAAATTTCAAAAATTTGGGCTGAAATCCAT 2715
QY 839 ACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGACCAAGTGGCGCGAAGCTGGTGG 898
DB 2716 ACATACTCCAGTATTTGCCATAAAGAAAAGACAGTACTAAATGGAGAAAATTAGTAG 2775
QY 899 ACTTCGGCGAGCTGAACAAGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC 958
DB 2776 ATTTTCAGAGAACTTAATAGAGAACTCAAGATTTCTGGGAAATTCATTAAGGAATACAC 2835
QY 959 ACCCGCGCGCTGAAGAGAGAGAGACGCTGACCGTCTGCAGCTGGCGACGCGCTACT 1018
DB 2836 ATCCTGCAAGGTTAAACAGAGAAAATCAGTAACAGTACTGATGTGGGCGATGCATATT 2895
QY 1019 TCAGCGTCCCTGACAGGAGCTTCGCAAGTACACCGCTTCACCATCCCGAGCATCA 1078
DB 2896 TTTTCAGTCCCTAGATAAAGACTTCAGAAAGTATACATTTACCAATACCTAGTATAA 2955
QY 1079 ACAACGAGACCCCGGATCGCTACCACTACAGCTGCTGCCCGAGGGCTGGAAGGCA 1138
DB 2956 ACAATGACACACAGGATTAGATATACATAGTACATGCTTCCACAGGATGGAAGGAT 3015
QY 1139 CCCCAGCATCTTCAGAGAGCATGACCAAGATCTCGGAGCCCTTCGCGCGCCGCAACC 1198
DB 3016 CACCAGCAATATTCAGTGTAGCATGACAAAAATCTTAGAGCCTTTTAGAAAAACAAAAATC 3075
QY 1199 CCGAGATCGTATACCA-----GGCCCCCTGTAGCTGGCGAGCACCTGGAGATCG 1252
DB 3076 CAGCATAGTCTATCTATCAATACATGATGATTTGTATGATGATCTGACTTAGAAATAG 3135
QY 1253 GCCAGCACCGGCCCAAGATCGAGAGCTGCGCAAGCACCTGCTGCGTGGGGCTTCACCA 1312
DB 3136 GGCAGCATAGAAAAATAGAGAACTGACAAACATCTGTTGAGGTGGGATTTACCA 3195
QY 1313 CCCCAGCAAGAACACACAGAGAGGAGCCCCCTTCCTGTGATGGCTACAGCTGCACC 1372
DB 3196 CACCAGCAAAAAACATCAGAAAGAACCTCCATCTCTTGGATGGGTTATGAACCTCCATC 3255
QY 1373 CCGACAAGTGACCGCTGACGCCCTGAGCTGCGGAGAGGAGCTGGACCGTGAACG 1432
DB 3256 CTGATAATGACAGTACAGCCTTAGTCTGCGCAGAAAGAGCAGCTGACTGTCAATG 3315
QY 1433 ACATCCAGAACTGGTGGCAAGTGAACCTGGGCCAGCCAGATCTACCCCGGCATCAAG 1492
DB 3316 ACATACAGAAATTAGTGGGAAATTTGAATTTGGCAAGTCAGATTTATGCAGGGATTAAG 3375
QY 1493 TGGCCAGCTGTCAAGCTGTGCGCGCGGCGCAAGGCCCTGACGACATCTGCGCCCTGA 1552
DB 3376 TAAGCAATTAATGAACCTTTAGGGGAAACCAAGCACTAACAGAAAGTAGTACCCTAA 3435
QY 1553 CCGAGGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGGAGCCCGTGCACG 1612
DB 3436 CAGAAGAGCAGAGCTAGAACTGGCAGAAAACAGGGAGATTTCTAAAGAACCCGTACATG 3495
QY 1613 GCGTGTACTAGACCCCAAGAACGCTGTGGCGGAGATCCAGAAGCAGGGCCACGACC 1672
DB 3496 GAGTGTATTAGACCCATCAAAAGACTTAATAGCAGAAATACAGAAGCAGGGGCAAGGCC 3555
QY 1673 AGTGGACCTACCATCTACAGAGCCCTTCAGARCTCAAGACCGGCAAGTACGCCA 1732
DB 3556 AATGGACATATCAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGAAAAATATGAA 3615
QY 1733 AGATGGCGCACCCCAACCAAGCAGTGAAGCAGCTGACCGAGCGCGTGCAGAAATCG 1792
DB 3616 GAATGAAGGGTGGCCACACTAATGATGTGAACAATTAACAGAGCGAGTACAAAAATAG 3675
QY 1793 CCATGGAGAGCATCGTGTCTGGGCAAGACCCCAAGTTCCGCTGCCCCATCCAGAAGG 1852

DB 3676 CCACAGAAAGCATAGTAATAATGGGAAAGACTCCTAAATTTAAATTTACCCATACAAAGG 3735
QY 1853 AGACCTGGAGAGACTGGTGGACCGACTACTTGGCAGGCGCACCTGGATCCCCGAGTGGAGT 1912
DB 3736 AAACATGGGAAGCATGGTGGACAGAGTATTTGGCAAGCCACCTGGATTTCTGAGTGGAGT 3795
QY 1913 TCGTCAACACCCCCCTGGTGAAGCTGTGTACAGCTGAGAGAGGAGCCCATCATCG 1972
DB 3796 TTGTCATATACCCCTCCCTTAGTGAAGTTATGTTACAGTTAGAGAAAGAACCCCATANTAG 3855
QY 1973 GCGCGAGACCTTCTACGTGGACGGCGCCCAACCGCGAGACCAAGATCGGCAAGCGG 2032
DB 3856 GAGCAGAAACTTTCTATGTAGATGGCGAGCCAAATAGGGAACCTAAATTTAGGAAGACAG 3915
QY 2033 GCTAGCTGACCGACCGGGCGGCGAGAGATCGTGAGCCTGACCGAGAGCACCAACAGCA 2092
DB 3916 GATATGTAACGTACAGAGAGAGACAAAAAGTTGTCCCTTAACGGACACACAACAAATCAGA 3975
QY 2093 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGACGAGACGCGCAGCGAGTGAACATCG 2152
DB 3976 AGACTGAGTTTACAGCAATTCATCTAGCTTTGCAGGATTTAGAAAGTAAACATAG 4035
QY 2153 TGACCGACAGCGACTACGCCCTGGCGCATCATCCAGCGCCGAGCCGACAGAGCGGAGCG 2212
DB 4036 TGACAGACTCAAAATATGCATTTGGGAATTCACAGCACACCAAGATAGAGTGAATCAG 4095
QY 2213 AGCTGGTGAACCAAGATCATCGAGAGCTGATCAAGAAGAGAGAGTGTACTGTAGCTGGG 2272
DB 4096 AGTTAGTCAGTCAAAATAATAGAGCAGTTAAATAAAAAAGGAAAAAGTCTACCTGGCATGG 4155
QY 2273 TGCCCGCCCAAGGAGCTGCGCGCAAGCAGCAGCATCGACAAGCTGGTGACCAAGGCCA 2332
DB 4156 TACCAGCACAAAGAAATTTGGAGAAATGAACAAGTAGTGGTGTGCTGCTGCTGAGGAA 4215
QY 2333 TCCGGAAGTGTCTTCTCTGACGGCATCGAT 2364
DB 4216 TCAGSAAAGTACTATTTTATAGATGAATAGAT 4247

RESULT 7
US-08-388-353-1
; Sequence 1, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: McPhee, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388.353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9709 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-388-353-1

Query Match 44.7%; Score 1100.4; DB 3; Length 9709;
Best Local Similarity 68.0%; Pred. No. 1.7e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;

QY 14 TGCCGAGGCATGAGCCAGCCACCA---GCGCAACATCTCTGATGTCAGCGCAGCAACT 70
DB 1877 TGCTGAAGCAATGAGCCAAAGTAACAAATCCAGCTACCAATGATACAGAAAAGGCAATT 1936
QY 71 TCAAGGGCCCAAGCGCATCATCAAGTGTCAACTGCGGCAAGGAGGGCCACATCGCCC 130
DB 1937 TTAGGACCAAGAAAGACTGTTAAGTGTTCAAATTGTGCAAGAGGCGCATAGCCA 1996
QY 131 GCAACTGCGCGGCCCGCCGCAAGAAAGGGCTGCTGGAAGTGCAGGAGGGCCACACAGA 190
DB 1997 AAAATTGACGGGCGCTAGGAAAAGGGCTGTTGGAATGTGAAAAGGAAGGACACAAA 2056
QY 191 TGAAGGACTCACCGAGCGCCAGGCCAACTTCTCCGAGGAGACCTGGCCCTTCCCCCAGG 250
DB 2057 TGAAGAAATTGACTGAGAGACAGGCTAA-TTTTATGGAAGATCTGGCCCTTCCCACAG 2115
QY 251 GCAAGGCCCGCGAGTTCCCGAGCGAGCAGAAACCGCGCAACAGCCCAACAGCGCGGAGC 310
DB 2116 GGAAGGCCAGGAAATTCTTCAGAGCAGACAGAGCCACAGCCCAACAGAGAGAGC 2175
QY 311 TGAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGCGCGCGCGCGCGCA 364
DB 2176 TTCAGGTTTGGGAAGAGACAACAACTCCCTCTCAGAAGCAGGAGCGGATAGACAAGAA 2235
QY 365 -----CCCTGAATCTCCCGACATCACCTGTGGCAGCGCCCTCTGGTGAGGATCAAGG 418
DB 2236 CTGTATCTTTAGTTCCTTCAGATCACTCTTGGCAGCGCCCTCGTCACAAATAAGA 2295
QY 419 TGGCGCCAGATCAAGAGGCGCTGTGGACACCGCGCGCGCGCGCGCGCGCGCGCGGAGG 478
DB 2296 TAGGGGGCAATTAAAGGAAGCTCTATTAGATACAGAGCAGATGATACAGATTAGAAG 2355
QY 479 AGATGAGCCTGCGCGGCAAGTGGAAAGCCCAAGATGATCGCGGGGATCGCGGGCTTCATCA 538
DB 2356 AAATGAATTTGCCAGGAAGATGGAACCAAAAATGATAGGGGAAATTGGAGGTTTATCA 2415
QY 539 AGGTGCCAGTAGTACGACAGATCTGTATCAGATCTGCGCAAGAGGCGCATCGGACCG 598
DB 2416 AAGTAGACAGTAGATGATCAGATCTATAGAAATCTGCGGACATAAAGCTATAGGTACAG 2475
QY 599 TGCTGATGCGGCCACCGCGTGAACATCATCGCGCCCAACATGCTGACCCAGCTGGGCT 658
DB 2476 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTGACTCAGATTGGCT 2535
QY 659 GCACCTGAACTTCCCATCAGCCCCATCAGACCGGTGCGCGGTGAAGCTGAAGCCCGGCA 718
DB 2536 GCACCTTTAAATTTTCCCAATTAGTCTATTGAGACTGTACAGTAAATTTAAAGCCAGGAA 2595
QY 719 TGGACGGCCCCAAGGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
DB 2596 TGGATGGCCCCAAAGTTAAACAATTGGCCATTGACAGAAGAAAAAATAAAGCATATTAGTAG 2655
QY 779 CCATCTCGAGAGATGGAGAAGGAGGCAAGATCAACAAGATCGGCGCCCGCAGAACCCCT 838
DB 2656 AAATTGTACAGAAATGGAAGAGGAGAAATTTCAAAATTTGGCCCTGAAATCCAT 2715
QY 839 ACAACACCCCGTGTGGCCATCAAGAAAGGAGACAGCAACAAAGTGGCGCAAGCTGGTG 898

DB 2716 ACAATACTCCAGTATTGGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAAATTAGTAG 2775
QY 899 ACTTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCGAGCTGGGCATCCCC 958
DB 2776 ATTTAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCA 2835
QY 959 ACCCCGCGCGCCTGAAGAAGAAAGACGCTGACCGCTGGAGCTGGGGGACGCCCTACT 1018
DB 2836 ATCCTGCAGGTTAAACACAGAAAAATCAGTAACAGTACTGGATGTGGCGCATCATATT 2895
QY 1019 TCAGCTGCGCCCTGGAGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 1078
DB 2896 TTTCAAGTCTCCTTAGATAAAGACTTTCAGGAAGTATACCTGCAATTTACCTAGTATAA 2955
QY 1079 ACAACGAGACCCCGCATCCGCTACCACTACACAGTGTGCCCGCAGGCGTGGGAAGGCA 1138
DB 2956 ACAATGAGACACCGAGGATTAGATATCAGTACAATGTCTTCCACAGGATGGAAGAT 3015
QY 1139 GCCCAGCATCTTCCAGAGCAGCATGACCAAGATCTTGGAGCCCTTCCGGCGCCGCAACC 1198
DB 3016 CACCAGCAATATTCAGTGTAGCATGACAAAAATCTTAGAGCCTTTTGAAGAAACAAATC 3075
QY 1199 CCGAGATCGTGAATCACTA-----GGCCCCCTGTAGCTGGGCGAGCAGCTGGAGATCG 1252
DB 3076 CAGACATAGTCACTATCAATACATGATGATTTGTATGTAGGATCTGACTTAGAAATAG 3135
QY 1253 GCAGACACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGTGGGGCTTCCACCA 1312
DB 3136 GGAGCATAGAACAAAAATAGAGAACTGAGACAACATCTGTTAGGTGGGATTTACCA 3195
QY 1313 CCCCAGCAAGAACACCCAGAAAGAGCGCCCTTCTCTGTGATGGGTACGAGCTGCACC 1372
DB 3196 CACCAGAAAAAACATCAGAAAGAACCTCCATTCTTGGATGGGTTATGAACCTCATC 3255
QY 1373 CCGACAAAGTGGAGCTGCGACCCCATCGAGCTGCCGAGAGAGAGAGCTGACCCGTGAACG 1432
DB 3256 CTGATTAATGGACAGTACAGCCTATAGTGTCCAGAAAAGACAGCTGGACTGTCAATG 3315
QY 1433 ACATCCAGAAAGCTGGTGGCAAGCTGAACTGGGCGACGAGATCTACCCGGCATCAAGG 1492
DB 3316 ACATACAAATTAAGTGGGAAATTTGAATTTGGGCAAGTCAAGATTTATGCGAGGATTAAG 3375
QY 1493 TGGCCAGCTGTGCAAGCTGCTGCGCGCGCCCAAGCCCTGACCGACATCGTGCCCTGA 1552
DB 3376 TAAGGCAATTTATGTAACTTCTTAGGGCAACCAAAAGCACTAACAGAAAGTAGTACCCTAA 3435
QY 1553 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGAGATCTCTGCGGAGCCCGTGCACG 1612
DB 3436 CAGAAGAGCAGAGCTAGAACTGGCAGAAAACAGGGAGATTCTAAAAGAACCGGTACATG 3495
QY 1613 GCGTGTACTACGACCCCGCAAGGACCTGGTGGCGGAGATCCAGAAGCAGGCGCCAGCAC 1672
DB 3496 GAGTGTATTATGACCCATCAAAAGACTTATAGCAGAAATACAGAACAGGGGCAAGGCC 3555
QY 1673 AGTGGACCTACAGATCTACCGAGGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1732
DB 3556 AATGGACATATCAAAATTTATCAAGAGCCATTATAAAATCTGAAAACAGGAAATATGCAA 3615
QY 1733 AGATGCGACCGGCCACCAACAGACGTGAAGCAGCTGACCGGAGCGCGTGGAGAGATCG 1792
DB 3616 GAATGAAGGGTGGCCACACTAATGATGTAAACAATTTAAGCAGAGGAGTACAAAAATAG 3675
QY 1793 CCATGGAGACATCGTGTGGGCAAGACCCCAAGTTCCCGCTCCCATCCAGAGG 1852
DB 3676 CCACAAAGCATAGTAATATGGGAAAGACTCTCTAAATTTAAATTTACCATACAAAGG 3735
QY 1853 AGACTGGGAGACCTGGTGGACCGACTTACTGGCAGGCGACCTGGATCCCGAGTGGAGT 1912
DB 3736 AAACATGGGAAGCATGGTGGACAGAGTATTGGCAAGCCACTGGATTTCTTGGTGGGAGT 3795
QY 1913 TCGTGAACACCCCGCTCGTGAAGCTGTGGTACAGCTGGGAGAGAGGAGCCCATCATCG 1972

Db 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGTTACCAAGTTAGAGAAAGAACCCCAATAATAG 3855
QY 1973 GCGCCGAGACCTTCTAGTGGAGCGGCCCGCCACACCGGAGACCAAGATCGGCAAGCCG 2032
Db 3856 GAGCAGAACTTCTATGTAGTGGGCGCCCAATAGGGAACCTAAATTAGGAAAGCAG 3915
QY 2033 GCTACGTGACCGACCGGGCGGCGAGAAAGATCGTGAAGCTGACCGAGACCAACCAACAGA 2092
Db 3916 GATATGTAAGTACAGAGAGAGACAAAGATGTTCCCCCTAACGGACACACAAATCAGA 3975
QY 2093 AGACCGAGCTGACAGGCATCCAGCTGGCCCTGCAGGACACGCGGAGGAGGTGAACATCG 2152
Db 3976 AGACTGAGTTACAAGCAATTCATCTAGCTTTGCGAGATTGCGGATTAGAAGTAACATAG 4035
QY 2153 TGACCGACAGCCAGTACGCTTGGGCATCATCCAGGCGCCAGCCGACCAAGAGCGAGAGCG 2212
Db 4036 TGACAGACTCACAAATATGCAATGGGAATCATTCAGGACACCAACAGATAGAGTGAATCAG 4095
QY 2213 AGCTGTTGAACCAAGATCATCGAGCAGCTGATCAAGAAAGGAGAGGTGTACTGAGCTGGG 2272
Db 4096 AGTTAGTCAGTCAATATAGAGCAGTTAATAAAAGGAAAGTCTACCTGGCATGGG 4155
QY 2273 TGCCCGCCCAAGAGGCGATCGGGGCAACGAGCAGATCGACAGCTGGTGAAGAGGCA 2332
Db 4156 TACCAGCACACAAGGAATTTGGAGGAATGAACAAGTAGATGGTTCAGTCTGCTGGAA 4215
QY 2333 TCCGCAAGGTGCTGTTCTCTGGACGGCATCGAT 2364
Db 4216 TCAGGAAGTACTATTTTATAGTGAATAGAT 4247

RESULT 8

US-08-488-551B-1

; Sequence 1, Application US/08488551B
; Patent No. 6015661
; GENERAL INFORMATION:
; APPLICANT: Nicholas J. Deacon
; APPLICANT: Dale A. McPhee
; APPLICANT: David Cooper
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 841
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,551B
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PK3864 (AU)
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: PK4002 (AU)
; FILING DATE: 21-FEB-1994
; APPLICATION NUMBER: PK0284 (AU)
; FILING DATE: 23-DEC-1994
; APPLICATION NUMBER: US 08/388,353
; FILING DATE: 14-FEB-1995
; APPLICATION NUMBER: PK3021/95
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FRANK S. DIGILIO
; REFERENCE/DOCKET NUMBER: 9606Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9709 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-488-551B-1

Query Match 44.7%; Score 1100.4; DB 3; Length 9709;
Best Local Similarity 68.0%; Pred. No. 1.7e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;
QY 14 TGGCCGAGGCGCATCAGCGCCACCA---CGCGCAACATCCTGTATGTCGCGCAGCAACT 70
Db 1877 TGGCTGAAGCAATGAGCCAGTAACAATCCAGTACCATAATGATACAGAAAGCAATT 1936
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGGAGGCGGCACATCGGCC 130
Db 1937 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATTTGGCAAAAGAGGGCACAATAGCCA 1996
QY 131 GCAACTGCCCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACACAGA 190
Db 1997 AAAATTGCAAGGCGCCCTAGGAAAGAGGGCTGTTGAAATGTGAAAGAGGAGCACACAAA 2056
QY 191 TGAAGGACTGCACGAGCGCCGACCAACTTCTTCCGCGAGGACCTGGCCTTCCGCCAGG 250
Db 2057 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTATAGGGAAGATCTGGCCTTCCCAAG 2115
QY 251 GCAAGCGCCGCGAGTTCCCGAGCGAGCAGAACCGCGCCCAACAGCCCGCCAGCGCGAGC 310
Db 2116 GGAAGGCGAGGGAATTTTCTTCAGAGCAGACACAGCCCAACAGCCCGCCAGAGAGAGC 2175
QY 311 TGAAGTGGCGCG-----CGACAACCGCCGAGGCGCGCGCGCGCGCGCGCGAGGCA 364
Db 2176 TTCAGGTTGGGGAAGAGACAACTCCCTCTCAGAAAGAGGAGCGCATAGACAAGAA 2235
QY 365 -----CCCTCAACTTCCCGCAGATCACCTGTGGCAGCGCCCTGCTGTGAGCTCAAGG 418
Db 2236 CTGTATCGTTAGCTTCCCTCAGATCACTTTTGGCAGCGCCCTGCTCACAATAAGA 2295
QY 419 TGGCGGCGCAGATCAAGAGGCGCCCTGTGGACACCGCGCGCGCGAGACACCTGTGGAGG 478
Db 2296 TAGGGGGCAATTAAAGGAAGTCTTATAGATACAGGAGCAGATGATACAGATTAGAAG 2355
QY 479 AGATGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGGGGCGATCGCGGCTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGGAACCAAAATGATAGGGGGAATTTGAGGTTTATCA 2415
QY 539 AGGTGCGCAGTACGACCAAGATCCTGTATCGAGATCTGCGGCAAGAAAGCCATCGCACCG 598
Db 2416 AAGTAGGACAGTATGATCAGATCACTATAGAAATCTGGGACATAAAGCTATAGTAGAG 2475
QY 599 TGCTGATCGGCGCCACCGCCCGTGAACATCATCGGCGCGCAACATGCTGACCGAGTGGGCT 658
Db 2476 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGAACTTCCCGATCAGCCCGATCGAGACCGTCCCGCTGAGCTGAGCGCCGCGCA 718
Db 2536 GCACCTTAAATTTCCCGATTTAGTCTTATGAGACTGTACCAAGTAAATTAAGCCAGGAA 2595
QY 719 TGGACGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAAATCAAGGCCCTTGACCG 778
Db 2596 TGGATGGCCCAAGTAAACAATGGCCATTTGACAGAGAAATAAAGCAATTAGTAG 2655
QY 779 CCATCTGCGAGGAGATGGAGAGGAGGCAAGATCACCAAGATCGCCCGCGCAAGACCCCT 838
Db 2656 AAATTTGTACAGAAATGGAAGGAGGAAATTTTCAAAATTTGGGCTGAAATTCAT 2715
QY 839 ACACACCCCGTGTTCGCCATCAAGAGAGGAGCAGCACCAAGTGGCGCGAGCTGGTGG 898
Db 2716 ACAATCTCCAGTATTTGCCATAAAGAAAAAGACAGTAAATGATGAGAAAAATTAGTAG 2775

QY	899	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCC	958
Db	2776	ATTTTCAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACAC	2835
QY	959	ACCCCGCGCGCTTCAGAGAGAAGAGCGTGACCGTGTGGACGTGGCGACGCCCTACT	1018
Db	2836	ATCTCGCAGGGTTAAACACAGAAAAATCAGTAACAGTACTGATGTGGCGCATGCATATT	2895
QY	1019	TCAGCGTGCCTCGGACGAGGACTTCGCAAGTACACCGCTTCACCATCCCCAGCATCA	1078
Db	2896	TTTCAGTTCCTTAGATAAAGACTTCAGAGAGTATATGTCATTTACCATACCTAGTATAA	2955
QY	1079	ACAACGAGACCCCGGATCCGCTACCACTACAACGTCTGCCCCAGGGCTGGAAGGCCA	1138
Db	2956	ACAATGAGACACCAAGGATAGATATCAGTACAATGTGCTTCCACAGGGATGGAAGGAT	3015
QY	1139	GCCCCAGCATTTCCAGAGCAGCATGACCAGATCCTGGAGCCCTTCGCGCCCGCAACC	1198
Db	3016	CACCAGCAATATTCCAGTGTAGCATGACAAAATCTTAGAGCCCTTTAGAAAACAAAATC	3075
QY	1199	CCGAGATCGTGATCTACCA-----GGCCCCCTCTACGTGGCGAGCGACCTCGAGATCG	1252
Db	3076	CAGACATAGTCATCTATCAATCATCGATGATTTGTATGTAGATCTGACTTAGAAATAG	3135
QY	1253	GCCAGCACCGGCCAAGATCGAGGAGCTGCGCAACCACTGCTGCGTGGGGCTTCACCA	1312
Db	3136	GGCAGCATAGACAAAAATAGAGAACTGAGACAACATCTGTTGAGGTGGSGATTTACCA	3195
QY	1313	CCCCGACAGACACCAAGAGGAGCCCTTCTGTGTGATGGCTACGAGCTGCACC	1372
Db	3196	CACCAGACAAAACATCAGAAGAACCTTCCATTCTTTGGATGGGTATTAACATCCATC	3255
QY	1373	CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAAGAGGAGAGCTGAGACGTGAACG	1432
Db	3256	CTGATAATGGCAGTACAGCCTATAGTCTGCCAGAAAAGACAGCTGAGACTGCAATG	3315
QY	1433	ACATCCAGAAGCTGGTGGCAAGCTGAATGGCGCAGCCAGATCTACCCGGCATCAAGG	1492
Db	3316	ACATACAGAAATTAGTGGGAAAATTGAATTGGGCAAGTCAGATTTATCGCAGGATTTAAG	3375
QY	1493	TGCGCCAGCTGTCAAGCTGTCTGCGGGCGCCAGGCCCTGACCGACATGTGTGCCCTGA	1552
Db	3376	TAAGGCAATTTATGTAACTTCTTAGGGAAACCAAGCACTAACAGAAAGTAGTACCACATA	3435
QY	1553	CCGAGGAGCGCAGCTGGAGCTGGCGAGAACCGCGAGATCCTCGCGAGCGCGTGCACG	1612
Db	3436	CAGAGAAGCAGAGCTAGACTGGCAGAAAACAGGGAGATTCTTAAAGAACCGGTACATG	3495
QY	1613	CGGTGTACTACGACCCCAAGAGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACC	1672
Db	3496	GAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGNAGCAGGGCAAGGCC	3555
QY	1673	AGTGGACCTACAGATCTACCAAGGACCTTCAGAACCTGAAACACCGGCAAGTACGCCA	1732
Db	3556	AATGGACATATCAAAATTTATCAAGAGCCATTTAATAATCTGAAAACAGGAAATATGCAA	3615
QY	1733	AGATGGCCACCGCCACCAACACGACTGAAGCAGCTGACCGAGCGCGTGCAGAAGATCG	1792
Db	3616	GAATGAAGGTGCCACACTATGATGTGAACATTTAACAGAGCAGTACAAAAAATAG	3675
QY	1793	CCATGGAGAGCATGTGTATCTGGGCAAGACCCCAAGTTCCTCGGCTGCCCATCCAGAAAG	1852
Db	3676	CCAGAAAAGCATAGTAATATGGGAAAAGACTCCTAAATTTAAATTTACCCATACAAAAGG	3735
QY	1853	AGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGATCCCGAGTGGGAGT	1912
Db	3736	AAACATGGGAGAGCATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAGTGGGAGT	3795
QY	1913	TCGTGAACACCCCCCTTGGTGAAGCTGTGTACCAAGCTGGAGAAGGCCCATCATCG	1972
Db	3796	TTGTCAATACCCCTCCCTTAGTGAAGTTATGGTACCAGTTAGAGAAAGAACCCATATAG	3855
QY	1973	GCGCCGAGACCTTCTACGTGGAGCGCGCCGCAACCCGCGAGACCAAGATCGGCAAGGCCG	2032

```

RESULT 9
US-09-309-572-15
; Sequence 15, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 9709
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-309-572-15

```

Query Match 44.7%; Score 1100.4; DB 4; Length 9709;
Best Local Similarity 68.0%; Pred. No. 1.7e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;

Db 2176 TTTAGGTTGGGGAAGAGACAACAACCTCCCTCTCAGAAGCAGGAGCCGATAGACAAGGAA 2235
QY 365 -----CCCTGNACTTCCCCAGATCACCCCTGTGCGACGCGCCCTGTGTAGCATCAAGG 418
Db 2236 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGGACGCGACCCCTCGTCAACAATAAGA 2295
QY 419 TGGCGGCGCAGATCAAGAGGCGCTGTGGACACCGCGCGGACACACCGTGTGGAGG 478
Db 2296 TAGGGGGCAATTAAGGAGAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAG 2355
QY 479 AGATGAGCCTGCCCGGCAAGTGAAGCCCAAGATGATCGGGGCGATCGCGGCGCTTCATCA 538
Db 2356 AAATGAATTTGCGAGGAAGATGGAACCAAAAATGATAGGGGGAATTTGAGGTTTTATCA 2415
QY 539 AGGTGCGCAGTAGCAGCAGATCTTGATCGAGATCTGCGGCAAGAGGCCATCGCACCG 598
Db 2416 AAGTAGGACATGATGATCAGATACTCATAGAAATCTGCGGACATAAAGCTATAGGTACAG 2475
QY 599 TGCTGATCGGCGCCACCCCGTGAACATCATCGGCGGCAAGATGCTGACCCAGCTGGGCT 658
Db 2476 TATTAGTAGGACCTACACCTGTACACATAATTGGAGAAATCTGTGTACTCAGATTGGCT 2535
QY 659 GCACCTTAAATTTCCCATTTAGTCTTGTAGACTGTACCAGTATAAAGTAAATTAAGGCCAGGA 2595
Db 2536 TGGAGGCGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
QY 719 TGGAGGCGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 2596 TGGATGGCGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 2655
QY 779 CCATCTCGGAGGAGATGGAGAAGGAGGCAAGATCACCAGATCGGCGCCCGAGAACCCCT 838
Db 2656 AAATTTGTACAGAAATGAAGAGGAGGAAATTTTCAAAATTTGGGCTGAAATCCAT 2715
QY 839 ACAACACCCCGTGTTCGCCATCAAGAGAAGGAGAGACAGCACCAGTGGCGCAAGCTGGTGG 898
Db 2716 ACAATACTCCAGTATTTGCCATTAAGAAAAGAGACAGTACTAAATGGAGAAAATTAGTAG 2775
QY 899 ACTTCCGCGAGCTGAACAGGCGACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCC 958
Db 2776 ATTTCAGAGAACTTAATAGAGAACTCAAGATTTCTGGGAGTTCAATTAGGAATACCAC 2835
QY 959 ACCCGCGCGGCTGAAGAGAAGAGAGCGTGACCGTGTGACGTGGCGCAGCGCTACT 1018
Db 2836 ATCCTGCGAGGTTAAACAGAAAAATCAGTAACAGTACTGGATGTGGCGATGATATT 2895
QY 1019 TCAGCGTCCCTTGAACGAGGACTTCCGCAAGTACACCGCTTACCATGCCAGCATCA 1078
Db 2896 TTTCAAGTTCCCTTAGATAAAGACTTCAGGAAGTATACCTATACCTAGTATAA 2955
QY 1079 ACAACGAGACCCCGGCATCCGCTACCACTACAGCTGCTCCCGCAGGCTGGAAGGCA 1138
Db 2956 ACAATGAGACACACAGGAAATAGATATCAGTACAATGTCTCCACAGGGATGGAAGGAT 3015
QY 1139 GCCCGAGCATCTTCAGAGCAGTACGACCAAGATCCTGGAGCCCTTCGCGCGCCGCAACC 1198
Db 3016 CACCAGCAATATTCCAGTGTAGCATGACAAAACTTTAGAGCCCTTTAGAAAAACAATC 3075
QY 1199 CCGAGATCGTGATCPACCA-----GGCCCCCTGTAGCTGGGCGAGCACCTGGAGATCG 1252
Db 3076 CAGACATAGTCATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAAAATAG 3135
QY 1253 GCCAGCAGCGCGCCAGATCAGGAGCTGCGCAAGCAGCTGCTGGCTGGGCTTCACCA 1312
Db 3136 GCGCATAGAACAAAAATAGAGGAACCTGAGACAACTCTGTGTAGGTGGGATTTAGCA 3195
QY 1313 CCCCAGCAAGAACCAAGCAGGAGGAGCCCTTCTGTGTGATGGCTACGAGCTGCACC 1372
Db 3196 CACCAGACAAAAACATCAGAAAGAACCTCCATCTCTTTGATGGGTTATGAACTCCATC 3255
QY 1373 CCGACAAGTGGACCGTGCAGGCCATCGAGCTGCCGAGAAAGAGAGCTGGACCGTGAACG 1432

Db 3256 CTGTAAATTTGGACAGTACAGCCCTATAGTGTGCCAAAAAGACAGCTGGAAGTGTCAATG 3315
QY 1433 ACATCCAGAGCTGTGGGCAAGCTGAAGTGGCCAGCCAGAGATCTACCCCGCATCAAGG 1492
Db 3316 ACATACAGAAATTAGTGGGAAAATTAATTTGGCAAGTCAGATTTATGCAGGATTAAG 3375
QY 1493 TCGCGCAGCTGTGCAAGCTGTGCGCGGCGCAAGGCCCTGACCCAGCATCTGTGCCCTCA 1552
Db 3376 TAAAGCAATTAATTAACCTTTTAGGGGAACCAAGCACTAACAGAACTAGTACCCTAA 3435
QY 1553 CCGAGAGCGCGAGCTGGAGCTGGCGCGAGAACCGGAGAGATCTGCGCGAGCCCGTGCAGC 1612
Db 3436 CAGAAGAAGCAGAGCTAGAACTGGCAGAAACACAGGAGATTTCTAAAAGAACCCGGTACATG 3495
QY 1613 GCGTGTACTACGACCCAGCAGCAGGAGCTGTGCGCGAGATCCAGAGAGCGGCCAGCAC 1672
Db 3496 GAGTGTATTATGACCCATCAAAAGACTTAATAGAGAAATACAGAGAGGGGCAAGGCC 3555
QY 1673 AGTGACCTTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGCCGCGCAAGTACGCCA 1732
Db 3556 AATGACATATCAATTTATCAAGAGCCATTTAANAATCTGAAACAGGAAATATGCAA 3615
QY 1733 AGATGGCAGCCGCCACACACAGCAGCTGAAGCAGCTGACCGAGGCGCTGCAGAAAGATCG 1792
Db 3616 GAATGAAGGGTCCCCACACTAATGATGTGAAACAATTAACAGAGGCGAGTACAAAAATAG 3675
QY 1793 CCATGGAGAGCATCTGATCTGGGCGACAGACCCCAAGTTCCGCTGCCCATCCAGAAAG 1852
Db 3676 CCACAGAAGCATAGTAATATGGGAAAGACTCTTAATTTAAATTTACCCATACAAAAGG 3735
QY 1853 AGACCTGGGAGACCTTGTGGACCGCTACTGTGCGAGCCACTGGATCCCGAGTGGGAGT 1912
Db 3736 AAACATGGGAACCATGTTGGACAGAGTATTGGCAAGCCACTGGATTTCTGTAGTGGAGT 3795
QY 1913 TCGTGAACACCCCGCTGTGAAAGCTGTGTACAGCTGTGAGAGAGGAGCCCATCATCG 1972
Db 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGTACAGTTAGAGAAAGAACCCATATAG 3855
QY 1973 GCGCGGAGACCTTCTACGTGGAGCGCGCGCAACCGCAGAGACCCAGATCGCAAGCCG 2032
Db 3856 GAGCAGAACTTTCTATGTAGATGGGCGCAATAGGGAACCTAAATTAGGAAAAGCAG 3915
QY 2033 GCTAGTGACCCGACCGCGCGGCGGAGAGATCTGTAGCCTGACCGAGACCCACCAACGAG 2092
Db 3916 GATATGTAATCAGAGAGGAAGACAAAAGTTGTCCCCCTAACGGACACACAAATCAGA 3975
QY 2093 AGACCGAGCTGCAGGCGATCCAGCTGGCCCTGCAGAGCAGCGGCGAGGAGTGAACATCG 2152
Db 3976 AGACTGAGTTTACAAGCAATTCATCTAGCTTTCAGGATTCGGGATTTAGAAGTAAACATAG 4035
QY 2153 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCCGACAGAGCGGAGCGG 2212
Db 4036 TGACAGACTTCAATATGCAATTTGGGAATCATTTCAAGCACAACCCAGATAAGAGTGAATCAG 4095
QY 2213 AGCTGGTGAACCCAGATCATCGAGCAGCTGATCAAGAAGAGAGAGGTGACCTGAGCTGGG 2272
Db 4096 AGTTAGTCAGTCAATATATAGAGCAGTAAATAAAGAGAAAGTCTACCTACCTGGCATGG 4155
QY 2273 TGCCCGCCCAAGGCGATCGCGGCGACAGCAGCATCGCAAGCTGTGTGAGCAAGGCA 2332
Db 4156 TACCAGCACACAAGGAATTTGGAGGAAATGAACAAGTAGATGGGTTGGTGTGCTGTGGAA 4215
QY 2333 TCCGCAAGGTGTCTCTCTGGACGCGCATCGAT 2364
Db 4216 TCAGAAAAGTACTATTTTATAGATGAATAGAT 4247

RESULT 10

US-08-935-312-13

; Sequence 13, Application US/08935312

; Patent No. 6207455

; GENERAL INFORMATION:

; APPLICANT: CHANG, Lung-Ji

;; TITLE OF INVENTION: LENTIVIRAL VECTORS
;; NUMBER OF SEQUENCES: 20
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BROWDY AND NEWMARK, P.L.L.C.
;; STREET: 624 Ninth Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20001
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/935,312
;; FILING DATE: 22-SEP-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: COOPER, Iver P.
;; REGISTRATION NUMBER: 28,005
;; REFERENCE/DOCKET NUMBER: CHANG=112
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 12494 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; DESCRIPTION: /desc = "DNA"
US-08-935-312-13

Query Match 44.7%; Score 1100.4; DB 4; Length 12494;

Best Local Similarity 68.0%; Pred. No. 1.8e-172;

Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;

QY	14	TGCGGAGGCGCATGAGCGAGGCGCACCA---GCGCCAACTCTCTGATGCGAGCGCAACT	70
DB	1548	TGCGTGAAGCAATGAGCCAGTAACTACCACTACCTACCTACCTACCTACCTACCTACCT	1607
QY	71	TCAAGGCGCCCAAGCGGATCATCAAGTCTTCAACTGCGGCAAGGAGGCGGCACATCGCCC	130
DB	1608	TTAGGAACCAAGAAAGAACTGTTAAGTGTTCATTTGTCGCAAGGAGGCGGCACATAGCCA	1667
QY	131	GCAACTGCCCGCCCGCCGCAAGAGGCTGCTGGAGTGGCGCAAGGAGGCGGCACACAGA	190
DB	1668	AAATTTGCGAGGCGCCCTAGGAAAGAGGCTGTTGGAAATGTGGAAGGAGGACACCCAA	1727
QY	191	TGAAGGACTGCACCGAGCGCGGCGCAACTTCTTCGCGAGGAGCTGGCCCTTCCCGCAGG	250
DB	1728	TGAAGATTGTACTGAGACAGCGCTAA-TTTTATTAGGGAAGATCTGGCCCTTCCACAG	1786
QY	251	GCAAGGCGCGGAGTTCOCAGCGAGCAGACAAACCGCCCAACAGCCCGCAGCGCGAGC	310
DB	1787	GGAAGGCGAGGAATTTCTTCAGAGCAGACACAGAGCAGACAGCCCGCCAGAGAGAGC	1846
QY	311	TGCAGGTGCGCGG-----CGACACCCCGCAGCGAGGCGCGCGCGAGCGCCAGGGCA	364
DB	1847	TTCAGGTTGGGGAAGAGACAACTCCCTCTCAGAAAGCAGGAGCGGATAGACAAAGAA	1906
QY	365	-----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTTGGTGAGCATCAAGG	418
DB	1907	CTGTATCTTCTTACCTTCCCTCAGATCATCTTTTGGCAGCGACCCCTCTGTCACAAATAAG	1966
QY	419	TGGGCGCCAGATCAAGAGGCGCCCTGCTGGACACCGCGCGCGCGACACCGTCTGCGAGG	478
DB	1967	TAGGGGGCAATTAAGAGAGCTCTATTAGATACAGAGCAGATGATACAGTATTAGAAG	2026
QY	479	AGATAGCCTGCCCGCAAGTGGAAACCCCAAGATGATCGGGCGCATCGCGGGCTTTCATCA	538

DB	2027	AAATGAATTTGCCAGGAAGATGAAACCAAAATGATAGGGGGAATTTGGAGGTTTTATCA	2086
QY	539	AGGTGCGCCAGTACGACCGAGATCCTGATCGAGATCTCGGCAAGAGCCCATCGGCACCG	598
DB	2087	AGTAGGACAGTATGATCAGATCTATAGAAATCTCGGACATAAAGCTATAGGTACAG	2146
QY	599	TGCTGATCGGCCCCACCCCGTGAACATCATCGGCCCAACATGCTGACCCAGCTGGGCT	658
DB	2147	TATTAGTAGCACCTACACCTGTCAACATAATTTGGAAGAAATCTGTTGACTCAGATTGGCT	2206
QY	659	GCACCCTGAACCTTCCCGCATCAGCCCGCATCAGACCGCTGCGCGTGAAGCTGAAGCCCGCA	718
DB	2207	GCACCTTAAATTTTCCCATTTAGCTTATGAGACTGTACAGTAAATTAAGACCGAGAA	2266
QY	719	TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCG	778
DB	2267	TGGATGGCCCAAGGTTAAACATGGCCATTTGACAGAGAAATAAAGCATTAGTAG	2326
QY	779	CCATCTGCGAGAGATGGAGAAGGAGGCAAGATCAACCAAGATCGGCCCGCGAGAACCCCT	838
DB	2327	AAATTTGTACAGAAATGGAAGAGGAGGAAATTTTCAAAATTTGGGCTGAAATCCAT	2386
QY	839	ACAACACCCCGTGTTCGCCCATCAAGAGAGGACACACCAAGTGGCGCAAGCTGGTG	898
DB	2387	ACAATACTCCAGTATTTGCCATTAAGAAAGAAAGACAGTACTAAATGGAGAAATTTAGTAG	2446
QY	899	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCGAGTGGGATCCCCC	958
DB	2447	ATTTCAGAGACTTAAAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCAC	2506
QY	959	ACCCCGCGGCGCTGAAGAAAGAGCGTGACCGTGTGGACGTGGGCGGAGCGCTTACT	1018
DB	2507	ATCCTGCAGGTTTAAACACAGAAATTAAGTAACTAAGTAACTAAGTAACTAAGTAACT	2566
QY	1019	TCAGCTGCCCGCTGGACGAGGACTTCCGCAAGTACACCCGCTTCACCATCCCCAGCATCA	1078
DB	2567	TTTCAGTCTCCCTTAGATTAAGAACTTTCAGGAAGTATAGTGCATTTACCATACCTAGTATA	2626
QY	1079	ACAACAGACCCCGGATCCGCTACCAAGTACAAAGTGCCTGCCCGAGGCTGGGAAGGCA	1138
DB	2627	ACAATGAGACACCGGATTTAGATATCAGTACATGCTGTCTCCACAGGATGGAAGGAT	2686
QY	1139	GCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCGCAACC	1198
DB	2687	CACGAGCAATATTTCCAGTGTAGCATGACAAATCTTTAGAGCCCTTTAGAAAACAAATC	2746
QY	1199	CCGAGATCGTATCTACCA-----GCCCCCTGTAGCTGGGCGAGGACCTGGAGATCG	1252
DB	2747	CAGACATAGTCTATCTCAATACATGATGATTTGTATGTAGGATCTGACTTAGAAATAG	2806
QY	1253	GCCAGACCCCGCAGATCGAGGAGCTGCGCAAGCACTGCTGCGCTGGGCTTCAACCA	1312
DB	2807	GGCAGCATAGAACAAATAGAGAACTGAGACACATCTGTTGAGTGGGGATTTACCA	2866
QY	1313	CCCCGACAAAGACACCGAGAGGAGCCCGCTTCTCTGTGGATGGGCTACGAGCTGCACC	1372
DB	2867	CACCAGACAAACATCAAGAAAGAACCTCCATTCCTTTGTGATGGGTTATGAATCCATC	2926
QY	1373	CCGACAGTGGACCGTGCACCCCATCGAGCTGCCGAGAGGAGAGCTGCAACCGTGAACG	1432
DB	2927	CTGATAATGGACAGTACAGCTTATAGTGTGCCGAGAAAGGACAGCTGGACTGTCAATG	2986
QY	1433	ACATCAGAGCTGGTGGGCAAGCTGAACCTGGGCGAGCCAGATCTACCCCGGATCAAGG	1492
DB	2987	ACATACAGAAATTTAGTGGGAAATTTGATTTGGCAAGTCAAGTTTATGAGGATTTAAG	3046
QY	1493	TGCGCGCAGCTGTCAAGCTGCTGCGGGCGCCCAAGGCCCTGACCGACATCGTCCCGCTGA	1552
DB	3047	TAAGGCAATTTATGAACTCTTTAGGGGAAACCAAGACACTAACAGAAAGTAGTACCACATA	3106
QY	1553	CCGAGGAGCGCGAGCTGGAGCTGCGCAGAACCCGAGATCTCCGCGAGCCCGCTGCACG	1612
DB	3107	CAGAAAGACGAGAGTAGAACTGGCAGAAACAGGGGATTTCTAAAAGACCGGTCATGT	3166

ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 015280-249000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 1581 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: -
LOCATION: 1..15581
OTHER INFORMATION: /note= "pNlnSg11"
US-08-646-538-35

Query Match 44.7%; Score 1100.4; DB 3; Length 15581;
Best Local Similarity 68.0%; Pred. No. 1.8e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;

Qy 14 TGGCGAGGCCATGAGCCAGGCCACCA---GCGCCAAACATCCTGATGCGAGCGCAGCAACT 70
Db 1877 TGGCTGAAGCAATGAGCCAAAGTAACAAATCCAGCTACCCTATGATGATACAGAAAGGCAATT 1936
Qy 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGGAGGCGCACATCGGCC 130
Db 1937 TTAGAACCAAGAAAGACTGTTAAGTGTTCATATGTTGGCAAGAGGCGCACATAGCCA 1996
Qy 131 GCAACTTCGCGCGCCCCCGCGCAAGAGGCTGCTGGAAGTGGCGAAGGAGGCGCACACAGA 190
Db 1997 AAAATTGACGGCCCTTAGGAAAAAGGCTGTTGSAATGTGGAAGAGGAGCACACAAA 2056
Qy 191 TGAAGGACTGACCGAGCGCCAGGCAACTTCTTCCGGGAGGACTGCGCTTCCCGCCAGG 250
Db 2057 TGAAGATGTTGACTGAGAGACAGGCTAA-TTTTATAGGGAAGATCGGCCCTCCACAA 2115
Qy 251 GCAAGGCGCGGAGTTCGCCAGCGAGAGAACCGCGCCACAGCCGCCACACCGCCGAGC 310
Db 2116 GGAAGCCAGGGAATTTCTTCAGAGCAGACAGCCAGCCACAGCCGCCACAGAGAGAGC 2175
Qy 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGGCGCGCGCGCGCGAGGCA 364
Db 2176 TTCAGGTTTGGGGAAGAGACAAACAATCCCTCTCAGAGAGCAGGCGCGATAGACAGAA 2235
Qy 365 -----CCCTGAACCTTCCCGCAGATCACCTCTGTGGAGCGCCCTGTGTGAGCATCAAG 418
Db 2236 CTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAACAATAAGA 2295
Qy 419 TGGGCGGCGACATCAAGAGGCGCTGCTGGACACCGGCGCGCGAGCAGACCGTGTGGAGG 478
Db 2296 TAGGGGGCAATTAAGGAAGCTCTATTAGATACAGAGCAGATGATACATATTAGAG 2355
Qy 479 AGATGAGCCTGCCCGGAAGTGAAGCCCAAGATGATCGGGCGGATCGGCGGCTTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGGAACCAAAATGATAGGGGGAATTGGAGGTTTATCA 2415
Qy 539 AGGTGCGCCAGTACGACCATCTGATCGAGATCTCGGGCAAGAGCCATCGGACCG 598
Db 2416 AAGTAGACAGATGATGATGATCTATGATGATCTAGAAATCTCGGACATTAAGCTATAGGTACAG 2475
Qy 599 TGCATGTCGGCCCCCAGCCCGTGAACATCATCGGCCCGCAACATCTGACCCAGCTGGGCT 658
Db 2476 TATTAGTAGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2595
Qy 659 GCACCTTGAATTTCCCATCAGCCCCATCGAGACCGTGTCCCGTGAAGCTGAAGCCCGCA 718
Db 2536 GCACCTTTAAATTTTCCCATATTAGTCTATTGAGACTGTGACAGTAAATTAAGCCAGGAA 2595
Qy 719 TGGAGCGCCCAAGGTGAAGCAGTGGCCCTTGACCCGAGAGAGATCAAGGCCCTGACCG 778

Db 2596 TGGATGGCCCAAAAGATTAAACAATGGCCATTGACAGAGAAAAAATAAAGCATTAGTAG 2655
Qy 779 COATCTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAGATCGGCCCGGAGAACCCCT 838
Db 2656 AAATTTGTACAGAAATGGAAGAGGAAAAATTTCAAAATTTGGCGCTGAAAAATCCAT 2715
Qy 839 ACAACACCCCGTGTTCGCCATCAAGAGAAGACAGACCAAGTGGCCCAAGCTGGTGG 898
Db 2716 ACAATACTCCAGTATTGGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAAATTAGTAG 2775
Qy 899 ACTTCCGGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGACGCTGGGCATCCCC 958
Db 2776 ATTTTCAGAGAATTAATAAGAACTCAAGATTCTCTGGAAGTTCATTAATAGGAATACCAC 2835
Qy 959 ACCCGCGCGCTGAAGAGAAGAGAGGCTGACCGTGTGACGCTGGCGGACGCGCTACT 1018
Db 2836 ATCTTGCAGGGTTAAATCAGAAAAAATCAGTAACAGTACTGGATGTGGCGATGCATATT 2895
Qy 1019 TCAGCGTCCCTGACGAGGACTTCCGCAAGTACACCGCTTCCACCTCCCGCAGCATCA 1078
Db 2896 TTTCAAGTTCCCTTAGATAAAGACTTTCAGGAAGTATACTGCATTTACCATACCTAGTATAA 2955
Qy 1079 ACAACGAGACCCCGCGCATCCGCTACCACTACAAGTGTGCTCCCGAGGCTGGAAGGCA 1138
Db 2956 ACAATGACACACAGGATAGATATCAGTACAATGTGCTTCCACAGGATGGAAGGAT 3015
Qy 1139 GCCCGCAGATCTTCCAGAGCAGCATGACCAAGATCCTTGAGCCCTTCCGCGCCCGCAACC 1198
Db 3016 CACCAAGCAATATTCAGTGTAGCATGACAAAAATCTTAGAGCCTTTAGAAAAACAATC 3075
Qy 1199 CCGAGATCGTGATCTACCA-----GGCCCCCTCTAGTGGCGAGCGACCTGGAGTCG 1252
Db 3076 CAGACATGATCATCATCAATACATGATGATTTGTATGTAGGATCTGACTTAGAAATAG 3135
Qy 1253 GCCAGCACCGCGCCAAAGATCGAGGAGCTGCGCAAGCACCTGCTGGGCTGGGCTTCCACCA 1312
Db 3136 GGCAGCATAGAACAAAAATAGAGGAATGAGACAACATCTGTTGAGTGGGATTTACCA 3195
Qy 1313 CCCCCGCAAGAACGACCCAGAGGAGCGCCCTTCTGTGGATGGGCTACGAGCTGCGAC 1372
Db 3196 CACCAAGCAAAAAACATCAGAAAGAACCTCCATCTTTGGATGGGTTATGAACCTCATC 3255
Qy 1373 CCGACAAGTGGACCTGCGAGCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAAGC 1432
Db 3256 CTGATAATGGACAGTACAGCTTATGCTGCTCCAGAAAGAGCAGCTGGACTGTCAATG 3315
Qy 1433 ACATCCAGAAAGCTGTTGGGCAAGCTGAACTGGGCCAGCAGATCTACCCCGGCATCAAG 1492
Db 3316 ACATACAGAAATTAGTGGGAAAAATGAAATGGGCAAGTCAGATTTATGACGGGATTAAG 3375
Qy 1493 TGGCCGACCTGTCAAGCTGCTGCGCGGCGCAAGGCGCTGACCCAGATCTGTCGCCCTGA 1552
Db 3376 TAAGCAATATTGTAACCTTCTTAGGGGAACCAAGAGCTAACAGAAAGTAGTACCATAA 3435
Qy 1553 CCGAGGAGCGCGAGCTGGAGCTGGCCGAGAGACCGGAGATCTGCGGAGCGCGTGAAG 1612
Db 3436 CAGAAAGCAGAGCTAGAACTGGCAGAAACAGAGGAGATCTTAAAGAACCGGTPACATG 3495
Qy 1613 GCGTGTACTACGACCCCGAGCAAGGACCTGGTGGCGGAGATCCAGAAAGCAGGCGCACGACC 1672
Db 3496 GAGTGTATTAGCCCATCAAAAGACTTAATAGCAGAAATACAGAAAGCAGGCGCAAGGCC 3555
Qy 1673 AGTGACCTACAGATCTTACCAGGAGCCCTTCAAGAACTTCAAGACCGCTGAAGACCGCAAGTACGCA 1732
Db 3556 AATGACATATCAAAATTTATCAAGGCAATTTAAAAATCTGAAAAACAGGAAAAATGCAA 3615
Qy 1733 AGATGCGCACCGCCACACCAACGAGCTGAAGAGCTGACCGAGGCGCTGAGAGATCG 1792
Db 3616 GAATGAAGGTTGCCACACTAATGATGTAACATTTAACAGAGCGCATCAAAAAATAG 3675
Qy 1793 CCATGGAGAGCATCGTGATCTGGGCGAAGCCCCCAAGTTCGCGCTGCCCATCCAGAGG 1852

Db 3676 CCACGAAAGCATAGTAATATGGGAAAGACTCTCTAAATTTAAATTAACCATCAAAAAGG 3735
QY 1853 AGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGGAGTGGAGT 1912
Db 3736 AAACATGGGAAGCATGGTGACAGAGATATTTGGCAAGCCACCTGGATTCTCTAGTGGAGT 3795
QY 1913 TCGTGAACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGGCCCATCATCG 1972
Db 3796 TTGCTCAATACCTCCCTTAGTGAAGTATGGTACCAGTTAGAGAAAGAACCCATATAG 3855
QY 1973 GCGCCGAGACCTTCTACGTGGACGGCGCCGCAACCGCCGAGACCAAGATGGCAAGCGCG 2032
Db 3856 GAGCAGAAACTTTCTATGTAGATGGGCGAGCCCAATAGGGAATACTAAATAGGAAGAGCAG 3915
QY 2033 GCTACGTGACCGACCGGGCGGCGCAGAGATCGTGAGCCCTGACCGAGACCAACCAACAGA 2092
Db 3916 GATATTAAGTACAGAGGAGGAGCAAAAAGTTGTCCCTTAACGGACACACAAATCAGA 3975
QY 2093 AGACCGAGCTGACAGGCCATCCAGCTGGCCCTGCAGGACAGCGGAGGTGAACATCG 2152
Db 3976 AGACTGAGTTACAGCAATTCATCTAGCTTGCAGGATTCGGGATTAGAAGTAACATAG 4035
QY 2153 TGACGACAGCCAGTACGCCCTGGGATCATCCAGGCCCGCCGAGCCGACAGAGCGAGCG 2212
Db 4036 TGACAGACTCACAATATGCAATGGGAATCATTCAGGCACAAACAGATAAGAGTGAATCAG 4095
QY 2213 AGCTGGTGAACAGATCATCGACAGCTGATCAAGAGGAGAGGTGTACTGAGCTGGG 2272
Db 4096 AGTTAGTCAGTCAAAATAGAGCAGTAAATAAAAGGAAAAAGTCTACTGGCATGGG 4155
QY 2273 TGCCCGCCCAAGAGGCGATCGCGGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGCA 2332
Db 4156 TACGACGACAAAGGAAATGGAGGAAATGAACAAGTAGATGGTGGTGGTCTGCTGGAA 4215
QY 2333 TCCGCAAGTGTCTTCTCGGACGGCATCAT 2364
Db 4216 TCAGGAAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 13

US-09-503-222-35
; Sequence 35, Application US/09503222
; Patent No. 6265548
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Galtanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/503,222
; PRIOR APPLICATION DATA:
; CLASSIFICATION:
; APPLICATION NUMBER: US/08/646,538
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677

; REFERENCE/DOCKET NUMBER: 015280-249000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: 1..15581
; LOCATION: 1..15581
; OTHER INFORMATION: /note= "pNLSG11"
; US-09-503-222-35

Query Match 44.7%; Score 1100.4; DB 4; Length 15581;

Best Local Similarity 68.0%; Pred. No. 1.8e-172;

Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;

QY 14 TGGCCGAGGCCATGAGCCAGGCCACCA---CGGCCACATCTCTGATCGCAGCGCAGCAACT 70
Db 1877 TGGCTGAAGCAATGAGCCAAAGTAACAATCCAGTACCAATTAATGATACAGAAAGGCAATT 1936
QY 71 TCAAGGGCCCCAAGCGGCATCATCAAGTGTCTCAACTGCGCAAGGAGGCGCACATCGCCC 130
Db 1937 TTAGGAACCAAAAGAAAGACTGTTAAGTGTTCAAATTTGTGCAAGAGAGGCGACATAGCCA 1996
QY 131 GCAACTGCGCCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGCAGGAGAGGCGCACAGA 190
Db 1997 AAAATTGACGGCCCTTAGGAAAAAGGGCTGTTGGAATGTGGAAGAGAAAGGACACCAAA 2056
QY 191 TGAAGAGCTCAGCGAGCGCCAGGCCAACTTCTTCCCGAGGAGCTTGGCCCTTCCCCCAGG 250
Db 2057 TGAAGATTTGCTGAGAGACAGGCTAA-TTTTTTAGGGAAGATCTGGCCCTTCCCAAG 2115
QY 251 GCAAGCGCCGCGAGTTCCCGCAGCGAGCAAGAACCGCGCAACAGCCGCCACCGCCGCGAGC 310
Db 2116 GGAAGCCAGGGAATTTCTTCAGAGCAGACAGAGCCACAGCCGCCACCAAGAGAGAGC 2175
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGAGCGCGCAGGCA 364
Db 2176 TTCAGGTTTGGGGAAGAGACAACAACCTCCCTCTCAGAAGCAGGAGCGGATAGACAAGAA 2235
QY 365 -----CCCTGAACCTTCCCGCAGATCACCTGTGGCAGCGCCCTGTTGAGCATCAAG 418
Db 2236 CTGTATCTCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCTGTCACAATAAGA 2295
QY 419 TGGCGCGCCAGATCAAGAGGCGCTGCTGGACACCGCGCGCGAGCACCGTGTCTGGAGG 478
Db 2296 TAGGGGGCAATTAAAGCAAGCTCTATTAGATACAGAGCAGATGATACAGATTAGAG 2355
QY 479 AGATGAGCCTTCCCGGCAAGTGAAGCCCAAGATGATCGGCGGATCGGGCGCTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGGAACCAAAATATAGGGGAAATTTGAGGTTTTATCA 2415
QY 539 AGTGCCCGCAGTAGCAGCCAGATCCTGATCGAGATCTCGGCGCAAGAGCCATCGCACCG 598
Db 2416 AAGTAGGACAGATGATGATCAGATCTCATAGAAATCTCGGACATATAAGCTATAGGTACAG 2475
QY 599 TGCTGATCGGCGCCCGCCCGTGAACATCATCGGCCCAACATCTGTCAGCCAGCTGGGCT 658
Db 2476 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTTGAACCTTCCCATCAGCCCCATCGAGACCGTGGCCGCTGAAGCTGAAGCCCGCA 718
Db 2536 GCACCTTTAAATTTTCCCATTAGTCTCTATTGAGACTGTACCAAGTAAATTAAGGACAGAA 2595
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGCCCTGACCG 778
Db 2596 TGGATGCGCCCAAGGTTAAACATTTGGCCATTTGACAGAGAAAAAATAAAGCATTAGTAG 2655

QY	779	CCATCTGGCAGAGATGGAAAGAGGAGGCAAGATCACCAGATCGGCCCGCGAAGACCCCT	838
Db	2656	AAATTTGTACAAATGGAAAAGGAGGAAAAATTTCAAAAAATTTGGCCTCGAAAAATCCAT	2715
QY	839	ACAACACCCCGTGTTCGCCATCAAGAAGACAGACACCAAGTGGCCGCAAGCTGGTGG	898
Db	2716	ACNATCTCCAGTATTGGCATTAAGAAAAAGACAGTACTAAATGGAGAAAATTAGTAG	2775
QY	899	ACTTCCGGAGCTGAACAGCCACCCAGGACTTCTGGGAGTGCAGCTGGGCGCATCCCC	958
Db	2776	ATTTTCAGAGAACTTAATRAAGAACTCAAGATTCTCGGAAGTTCATTTAGGAATACCAC	2835
QY	959	ACCCCGCGCGCTGAAGAAGAAGAGCGTGACCGTGTGAGCGTGGGCGACGCGCTACT	1018
Db	2836	ATCTCTGCAGGTTAAACACAAAAAATCAGTAACTACTGGATGTGGCGATGCTATAT	2895
QY	1019	TCAGCGTGCCTTGGACGAGGACTTCCGCAAGTACACGCTTCACCATCCCAGCATCA	1078
Db	2896	TTTCAGTTCCCTTAGATTAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAA	2955
QY	1079	ACAACGAGACCCCGCGCATCCGCTACCAAGTACAAGTGTCTGCCACAGGCTTGAAGGGCA	1138
Db	2956	ACAATGAGACACAGGGATTAGATATCAGTACAATGTCTTCCACAGGATGGAAGGAT	3015
QY	1139	GCCCGCAGCATTTCCAGAGCAGCATGACCAAGATCCTGGAGCCTTTCGCGCCGCGCAAC	1198
Db	3016	CACCAAGCAATATTCAGTGTGATCATGACATGACAAAAATCTTAGAGCTTTTACAAAA	3075
QY	1199	CGAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGCAGCGACCTGGAGATCG	1252
Db	3076	CAGACATAGTCATCTATCAATACATGGATGATTGTATGAGATCTGACTTAGAANAATAG	3135
QY	1253	GCCAGCACCGGCCAAGATCGAGGAGCTGCGCAAGCACCTGTGCTGGCGTGGCGTTTACCA	1312
Db	3136	GSCAGCATAGAACAAAAATAGAGGAATGAGACACATCTGTTAGTGGGATTTACCA	3195
QY	1313	CCCCGCAAGAGCACACAGAGGAGCCCTTCTCTGTGATGGGCTACGAGCTGCACC	1372
Db	3196	CACCAAGAAAAACATCAGAAGAACCTTCCATTCCTTTGGATGGGTTATGAATCCATC	3255
QY	1373	CGCAAGTGGACCGTGCAGCCATCGAGCTGCCGAGAGGAGAGCTGGACCGTGAACG	1432
Db	3256	CTGATAATGGACGTACAGCCTATAGTGTGCCAGAAAGACAGCTGGATGTCAATG	3315
QY	1433	ACATCCAGAACTGTTGGGCAAGCTGAATGGGCCAGCCAGACATCTACCCCGCATCAAG	1492
Db	3316	ACATACAGAAATAGTGGGAAAATGAATTGGCAAGTCAGATTATCGAGGGATTAAG	3375
QY	1493	TGCGCCAGCTGTGAAGCTGTGCGCGGCCAAGCCCTGACCGACATCTGTCGCCCTGA	1552
Db	3376	TAAGCAATTTATGTAATCTTTAGGGAAACCAAGCACTAACAGAAAGTAGTACCACATA	3435
QY	1553	CCGAGGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGCGCGAGCCCGTGACG	1612
Db	3436	CAGAAGAGCAGCTAGACTTGGCAGAAAAACAGGAGATTCTAAAAGAACCGGTACATG	3495
QY	1613	GCCTGTACTAGACCCCAAGGACCTTGTGGCCGAGATCCAGAAGCAGAGGCCAGCAC	1672
Db	3496	GAGTGTATTATGACCCATCAAAAGACTTAATAGAGAAATACAGAAAGCGGCGAAGCC	3555
QY	1673	AGTGAACCTACAGATCTACAGGAGCCCTTCAAGAACTGGAAGACCGGCAAGTACGCCA	1732
Db	3556	AATGAAGGTTGCCACACTAATGATGTGAACAAATTAACAGAGCAGTACAAAAAATAG	3615
QY	1733	AGATCGCACCGCCACACCAACGACGTGAAGCAGCTGACCCAGGCGGTGCAGAGATCG	1792
Db	3616	GAATGAAGGTTGCCACACTAATGATGTGAACAAATTAACAGAGCAGTACAAAAAATAG	3675
QY	1793	CCATGGAGAGATCGTGATCTGGGCAAGACCCCAAGTTCCGCTGCCCTTCCAGAGG	1852
Db	3676	CCACAGAAGCATAGTAAATATGGGGAAGACTCTCTAAATTTAAATTTACCATACAAAG	3735
QY	1853	AGACCTGGGAGACCTTGTGGACCGCACTACTGGCAGCGCACCTTGGATCCCGAGTGGAGT	1912

Db	3736	AAACATGGGAAGCATGGTGGACAGAGATATTGGCAAGCCACCTGGATTCTCTAGTGGGAGT	3795
Qy	1913	TCGTGAACACCCCCCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCG	1972
Db	3796	TTGTCAATACCCCTCCCTTAGTGAAGTTATGGTACAGTTAGGAAGAAGAACCCATAATAG	3855
Qy	1973	GGCCGGAGACCTTCTACGTGGACGGCGCGCCCAACCGCAGACCAAGATCGGCNAGGCCG	2032
Db	3856	GAGCAGAACTTTCATGTATGTAGATGGCGACCCAATAGGGAACCTAAATTAGGAAAAGCAG	3915
Qy	2033	GCTACGTGACCGACCGGGCGCCGGCAGAGAAGATCGTGACCTGTACCGAGACACCAACGAGA	2092
Db	3916	GATATGTAACTGACAGAGGAAGACAAAAAGTTGTCCCTTACGGGACACAACAAATCAGA	3975
Qy	2093	AGACCGAGCTGCAGGCCATCCAGCTTGGCCCTCGAGGACAGCGGCACGAGGTGAACATCG	2152
Db	3976	AGACTGAGTTACAGCAATTCATCTAGCTTTCGAGGATTGAGANGTAAACATAG	4035
Qy	2153	TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCAGCCCGACAGAGCGAGCG	2212
Db	4036	TGACAGACTCAAAATATGCATTTGGGAATCATTCAGCACAAACCAGATAAGAGTGAATCAG	4095
Qy	2213	ASCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAAGAGAGAGGTGTACCTCAGCTGGG	2272
Db	4096	AGTTAGTCAGTCAAATAATAGAGCAGTTAATAAAAAAGGAAAAAGTCTACCTGGCATGGG	4155
Qy	2273	TGCCCGCCCAACAGGCGCATCGCGCGCAACGAGCAGATCGACAAGCTGGTGTAGCAAGGGCA	2332
Db	4156	TACCAACACAAAGGAATTTGGAGGAANAATGAACAAGTAGTAGTGGTGTGGTCTGGAA	4215
Qy	2333	TCCGCAAGTGTCTGTTCCTGGACGGCATCGAT	2364
Db	4216	TCAGAAAGTACTATTTTATAGATGGAATAGAT	4247

```

RESULT 14
US-08-944-449-7
; Sequence 7, Application US/08944449
; Patent No. 5985613
; GENERAL INFORMATION:
; APPLICANT: KURTH, REINHARD
; APPLICANT: BAIER, MICHAEL
; APPLICANT: METZNER, KARIN
; APPLICANT: WERNER, ALBRECHT
; TITLE OF INVENTION: Use of an "immunodeficiency-virus suppressing
; TITLE OF INVENTION: lymphokine (ISL)" to inhibit the replication of
; TITLE OF INVENTION: viruses, particularly of retroviruses
; FILE REFERENCE: 8341-7065
; CURRENT APPLICATION NUMBER: US/08/944,449
; CURRENT FILING DATE: 1997-10-06
; EARLIER APPLICATION NUMBER: EP 95113013.2
; EARLIER FILING DATE: 1995-08-18
; EARLIER APPLICATION NUMBER: DE 195 13 152.5
; EARLIER FILING DATE: 1995-04-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-08-944-449-7

```

	Query Match	44.3%	Score 1090.8	DB 2	Length 9737
Best Local Similarity	67.8%				
Matches 1608	Conservative	0	Mismatches 742	Indels 22	Gaps 5
Qy	14	TGGCGGAGCCATGAGCCAGGCCACCA---GGGCCAACATCCTGATGCAGCGCAGCAACT	70		
Db	1884	TGGCTGGAAGCCATGAGCCCAAGTAACAATCAGCTAAGCATATGATGCAGAGAGGCAATT	1943		
Qy	71	TCAGGGGCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGGCC	130		

Db	1944	TTAGGAACCAAAAGACGTGTTAAGTGTTCCTCAATTGTGGCAAAAGAGGGCAATAGCCA	2003
Qy	131	GCAACTGCGCGCCCCCGCCCAAGAAAGGCTGCTGGAAGTGCGGCAAGAGGGCCACCAGA	190
Db	2004	AAAATTGCAGGGCCCCTAGAAAAAGGCTGTTGGAGATGTGGAAGGAGGACACCAA	2063
Qy	191	TGAAGAGCTGCACCGAGCCGACGGCCAACTTCTTCGGCGAGGACCTGGCCTTCCCCGAGG	250
Db	2064	TGAAGAATTGCACTGAGACACAGGCTAA--TTTTTTTAGGGAAGATCTGGCCTTCTCTACAAG	2122
Qy	251	GCAGGCCGCGAGTTCCCCACGGAGCAGAACCGCGCCCAACAGACCCACACGCGCGAGC	310
Db	2123	GGAGGCCAGGGAAATTTCTTCAGAGCAGACAGAGCCAAACAGCCCCACCAAGAGAGC	2182
Qy	311	TGCAGGTGGCGGCGC-----ACAACCCCGCAGCGAGCGCGCGCGCGAGCGCCAGGCGCA	364
Db	2183	TTCAGGTTGGGAGGAGAAACACTCCCTCTCAGAGCAGGAGCGGTAGACACAGGAA	2242
Qy	365	-----CCCTGAACTTCCCCCAGATCAOCCCTGTGGCAGCGCCCGCTGGTGAAGCATCAAGG	418
Db	2243	CTGTATCTCTTAATCTTCCCTCAGATCACTCTTTGGCAACGACCCTCGTCACAATAAGGA	2302
Qy	419	TGGCGGCCAGATCAAGGAGGGCCCTGCTGGACACCGGCGCCGACGACACCGTCTGGAGG	478
Db	2303	TAGGGGGGCACTTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG	2362
Qy	479	AGATGAGCCTGCCCGCAAGTGGAAAGCCCAAGATGATCGCGGGATCGGGGCTTTATCA	538
Db	2363	AAATGAAATTTGCCAGGAAATGGAAACCAAAATGATAGGGGGAATTTGGAGGTTTATCA	2422
Qy	539	AGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAAGCCATCGGCACCG	598
Db	2423	AAGTAAGACAGTACGATCAGATACCTGTAGAAATCTGTGGACATAAAGCTATAGGTACAG	2482
Qy	599	TGCTGATGCGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCT	658
Db	2483	TATTAGTAGGACCTACACCTGTCAACATAAATTTGGAAGAAATCTGTTGACTCAGATTGGTT	2542
Qy	659	GCACCTGAACTTCCCCATCAGCCCATCGAGACCGTCCCGTGAAGCTGAAGCGCGGCA	718
Db	2543	GTACTTTAAATTTCCCCATTTAGTCTATTGAACTGTACAGTAAATTTAAGCCAGGAA	2602
Qy	719	TGGACGCGCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGATCAAAGCCCTGACCG	778
Db	2603	TGGATGCCCCAAAAGTTAAGCAATGCCATTGCACAGAAGAAAAATAAAAGCATTTAGTAG	2662
Qy	779	CCATCTCGGAGGAGATGGAGAGGAGGGCAAGATCACCAAGTCGGCCCCGAGNACCCT	838
Db	2663	AGATATGTACAGAAATTTGAAAAAGGAAGGAAATTTCAAAAATTTGGCCCTGAAATCCAT	2722
Qy	839	ACAACACCCCGTTTCGCCCATCAAGAGAGGAGCAGCACCAAGTGGCGCAAGCTGTGG	898
Db	2723	ACAATCTCCAGTATTTGCTATTAAGAAANAAGACAGTACTAAATGGAGAAACTAGTAG	2782
Qy	899	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC	958
Db	2783	ATTTACAGAGAATTAATAAAGAACTCAAGACTTCTGGGAAGTTCAGTTAGGAATACCAC	2842
Qy	959	ACCCGCGGCCCTGAAGAAGAGAGAGCGTGACCGTCTGGACGTGGGCGACGCTTACT	1018
Db	2843	ACCCGCGAGGGTTTAAAAAAGAAAAAATCAGTAACAGTATTGGATGTGGGTGATGCATCT	2902
Qy	1019	TCAGCGTGCCCTGGACGAGGACTTCCGCAAGTACACCGCCTTACCATCCCCAGCATCA	1078
Db	2903	TTTCAGTTCCTTTAGATAAAGACTTTAGAAAGTATPACTGTCATTTACCATACCTAGTATAA	2962
Qy	1079	ACAACGAGACCCCGGCATCCCGTACCAGTACCAAGTCTGCCCCAGGGCTTGAAGGGCA	1138
Db	2963	ACAATGAGACACAGGATTAGATATCAGTACATGTCTGCCACAGGGATGGAAGGAT	3022
Qy	1139	GCCCCAGATCTTTCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGGCCCGCAACC	1198
Db	3023	CACGACGAATATTCCAAAGTAGCATGACAAAATCTTAGAGCCCTTTTAGAAAACAGAACT	3082

QY	1199	CCGAGATCGTGAATCTACCA-----GGCCCCCTGTACGTGGCAGCGACCTGGAGATCG	1252
DB	3083	CAGACATAGTTATCTATCAATACATGGATGATTGTGTATAGGATCTGACTTTAGAANAATAG	3142
QY	1253	GCACGACCGCCCAAGATCGAGGAGCTCGCAAGCACCTGCTGCCTGGGCTCTCCACA	1312
DB	3143	GCACGATAGAACAAAATAGAGAACTGAGACAGATCTGTTGAGTGGGATTTTACCA	3202
QY	1313	CCCCGCAAGAAGCACCAAGAAGGAGCCCCCTTCCTGTGGATGGCTACGAGCTGCACC	1372
DB	3203	CACCAGACAAAACATCAGAAAGAACCTCCATTCTCTGGTGGGTATGAACTCCATC	3262
QY	1373	CGACAAAGTGGACGTCGAGCCCCATCGACCTGCCCGAAGAGAGACTGCACCGTGAACG	1432
DB	3263	CTGATAAATGGACGTACAGCCTATTAATCTGCCGAAAAGACACTGGACTGTCAATG	3322
QY	1433	ACATCCAGAAGCTGGTGGGCAAGCTCAACTGGGCCAGCCAGATCTACCCGGGCATCAAGG	1492
DB	3323	ACATACAGAAGTTAGTGGGAAATTCGAATGGCAGTCTCAGATTTATGCAGGGATTTAAG	3382
QY	1493	TGCGCCAGCTGTGCAAGCTGTGCGGGCGCAAGCCCTTGACCGACATCTGTGCCCTGA	1552
DB	3383	TAAAGCAGTTATGTAAACTCTTAGAGGAACCAAGACACTAACAGAAGTAATACCACATA	3442
QY	1553	CGAGGAGCCGAGCTGGAGCTGGCGGAGAACCGCAGAGATCCTGCGCGAGCCCGTGCACG	1612
DB	3443	CAGAAGACGAGCTAGAACTTGGCAGAAACAGGAGATCTTAAAGAACCTAGTACATG	3502
QY	1613	GGGTGTACTACGACCCCAAGAGACTCTGGTGGCGAGATCCAGAAAGAGGGCCAGGACC	1672
DB	3503	AAGTATATTACCCCATCAAAAGACTTAGTAGCAGAAATACAGAAGCAGGGGCAAGGCC	3562
QY	1673	AGTGACCTACCAAGTCTACAGGAGGCCCTTCAGAAACCTGAAGACCGCAAGTAGGCCA	1732
DB	3563	AATGGACATATCAATTTATCAAGAGCCATTTAAAAATCTGAAACAGGAAAGATATGCAA	3622
QY	1733	AGATCGGACCGCCACACCAACGAGCTGAAGCAGCTGACCAGGCGCGTGCAAGAAATCG	1792
DB	3623	GGATGGGGTCCCCACACTAATGATGTAACAGTTAACAGAGGCAGTGCAAAAAGTAT	3682
QY	1793	CGATGAGAGCATCTGTGATCTGGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAAG	1852
DB	3683	CCACAGAAGCATAGTAATATCGGGAAGATTCTTAATTTAACTACCCATACAAAAGG	3742
QY	1853	AGACCTGGGAGACCTGTGGACCGCACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGT	1912
DB	3743	AAACATGGGAACATGGTGGATGGAGTTTGCAACCTACCTGGATTCTGTAGTGGGAGT	3802
QY	1913	TCGTGAACACCCCCCCTGGTGAAGCTGTGTACAGCTGGGAAGAGGCCCATCATCG	1972
DB	3803	TTGTCAATACCCCTCCCTTAGTGAATATTAGTTACCACTTAGAGAAAGAACCCATAGTAG	3862
QY	1973	GGCGCGAGACCTTCTACGTGGAGCGCGCCCAACCGCAGACCAAGATCGGCAAGCCCG	2032
DB	3863	GAGCAGAACTTCTATGTAGATGGGGCAGCTAATAGGAGACTAATTAGGAAAAGCAG	3922
QY	2033	GCTAGCTGACCGACCGGGCCGGCAGAAGATCGTAGCCTGACCGAGACCCCAACACAGA	2092
DB	3923	GATATGTTACTGACAGAGGAACAAAAAAGTTGTCTCCATAGCTGACACACAACAAATCAGA	3982
QY	2093	AGACCGAGCTGCAGGCCATCCAGCTGGCCCCTCGAGACAGCGGCAGGAGTGAACATCG	2152
DB	3983	AGACTGAATTACAAGCAATTTCATAGCTTTCAGGATTCGGGATTTAGAGTTAAACATAG	4042
QY	2153	TCACCGACAGCCAGTACGCCCTTGGGCATCATCCAGGCCAGCCCGCAAGAGCGGACGCG	2212
DB	4043	TAAACAGCTCACAATATGCATTAGGAACTCATTCACGACACACCAAGATAGAGTGAATCAG	4102
QY	2213	AGCTGGTGAACCAAGTATCTGAGCAGCTGATCAAGAAGGAGAGGTTGTAACCTGAGCTGGG	2272
DB	4103	AGTTAGTCAGTCAAAATATAGAGCAGTTAATAAAAAGGAAAGAGTCTACCTGGCATGGG	4162

QY 2273 TGCCCCCACAAGGGCATCGCGCGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCA 2332
 Db 4163 TACCACACACAAGGAATGGAGGAAATGAACAAGTAGATAAATTAGTCAGTCTGGAA 4222
 QY 2333 TCAGCAGGTGCTGTTCTCTGGACGGCATCGAT 2364
 Db 4223 TCAGAAAGTACTATTTTGAATGGAATAGAT 4254

RESULT 15

US-09-353-362-7
 : Sequence 7, Application US/09353362
 : Patent No. 638739
 : GENERAL INFORMATION:
 : APPLICANT:
 : TITLE OF INVENTION: Use of an "immunodeficiency-virus suppressing
 : lymphokine (ISL)" to inhibit the replication of viruses,
 : TITLE OF INVENTION: in particular of retroviruses
 : NUMBER OF SEQUENCES: 8
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.30B (EPO)
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/353,362
 : FILING DATE: 15-JUL-1999
 : CLASSIFICATION:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: DE 195 13 152.5
 : FILING DATE: 07-APR-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: EP 95113013.2
 : FILING DATE: 18-AUG-1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: KLESNER, Sharon N.
 : REGISTRATION NUMBER: 36,335
 : REFERENCE/DOCKET NUMBER: P8341-9012
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202) 638-5000
 : TELEFAX: (202) 638-4810
 : TELEX:
 : INFORMATION FOR SEQ ID NO: 7:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 9737 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : MOLECULE TYPE: DNA (genomic)
 : US-09-353-362-7

Query Match 44.3%; Score 1090.8; DB 4; Length 9737;
 Best Local Similarity 67.8%; Pred. No. 6.5e-171;
 Matches 1608; Conservative 0; Mismatches 742; Indels 22; Gaps 5;
 QY 14 TGCGCCAGGCCATGAGCCAGGCCACCA---GCGCCAACATCTCTGATCGAGCGCAGCAACT 70
 Db 1884 TGGCTGAAGCCATGAGCCAGTAACAAATCCAGCTAACATAATGATCGAGAGGCAATT 1943
 QY 71 TCAAGGCCCCACGCGCATCATCAAGTGTCTCAACTCGCGGCAAGGAGGCCACATCGGCC 130
 Db 1944 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATATTGGCAAGAAAGGGGCACATAGCCA 2003
 QY 131 GCAACTCGCGCCCCCGCGAAGAGGCTGCTGGAAGTGGCGCAAGAGGCCACCCAGA 190
 Db 2004 AAAATTGCAGGCCCTTAGGAAAAGGGCTGTTGGAGATGTGGAAGGAGGACACCAAA 2063
 QY 191 TGAAGAGTCTGACCGAGCGCCAGCCCAACTTCTTCGCGAGGACTTGGCCTTCCGCCAGG 250
 Db 2064 TGAAGATTGCACCTGAGAGACAGGCTAA-TTTTATGGGAGATCTGGCCTTCTACAG 2122
 QY 251 GCAAGGCCCGGAGTTCCCGACGAGCAGAACCGCGCCACACAGCCCAACCGCGCGAGC 310
 Db 3203 CACCAGACAAAAACATCAGAAAGAACCTCCATTCTTTGGATGGGTTATGAACCTCCATC 3262

Db 2123 GGAAGGCCAGGGAATTTTCTTCAGAGCAGACGAGCCACAGCGCCACCACCAAGAGAGC 2182
 QY 311 TGCAGGTGCGGGCG- - - - -ACAACCCCGCAGCGAGGCGCGCGCGAGCGCGAGGCA 364
 Db 2183 TTCAGGTTTGGGAGGAGAAACAACTCCCTCTCAGAAGCAGGAGCGCATAGACAAGGAA 2242
 QY 365 - - - - -CCCTGAACCTTCCCCAGATCACCTTGTGCGACGCGCCCTGGTGAGCATCAAGG 418
 Db 2243 CTGTATCCTTTAACTTCCCTCAGATCACTTTTGGCAACGACCCCTCGTCACAATAAGGA 2302
 QY 419 TGGCGGCCAGATCAAGAGGCCCCCTGTGGACACCGGCGCGCGAGACACCTGCTGGAGG 478
 Db 2303 TAGGGGGCAACTAAGAGAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAG 2362
 QY 479 AGATGAGCCTCCCGCGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGGCTTCATCA 538
 Db 2363 AAATGAATTTGCCAGGAAATGGAAACCAAAATGATAGGGGAAATTGGAGTTTATCA 2422
 QY 539 AGGTGCGCCAGTAGCACCCAGATCCTGTGATCGAGATCTGCGGCAAGAGCCCATCGCACCG 598
 Db 2423 AAGTAAGACAGTACGATCAGATACCTGTAGAAATCTGTGGACATAAAGCTATAGTACAG 2482
 QY 599 TGCTGATCGGCCACCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT 658
 Db 2483 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGTT 2542
 QY 659 GCACCTGAACCTTCCCATCAGCCCATCGAGACGCTGCGCGTGAAGCTGAAGCCCGGCA 718
 Db 2543 GTACTTTAAATTTCCCATTTAGTCTTGAACCTGTACCAGTAAATTAAGCCAGGAA 2602
 QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAAATCAAGCCCTGACCG 778
 Db 2603 TGGATGCGCCAAAGTTAAGCAATGGCCATTGACAGAGAAATAAAGCATTAGTAG 2662
 QY 779 CCATCTGGGAGGATGGAGAAGGAGGCAAGATCACCAAGATCGGCGCCCGAGAACCCCT 838
 Db 2663 AGATATGTACAGAAATGGAAAGGAGGAAATTTCAAAATTTGGGCTGAAATCCAT 2722
 QY 839 ACAACACCCCGTGTGCGCATCAAGAAGAGGACAGACCAAGTGGCGCGAGCTGGTG 898
 Db 2723 ACAATCTCCAGTATTGCTATAAAGAAAAGACAGTACTAATGGAGAAACTAGTAG 2782
 QY 899 ACTTCCGCGAGCTGAACAAGCCACCCAGGACTTCTTGGAGGTGCGAGCTGGGCACTCCCC 958
 Db 2783 ATTTTCAGAGAATTAATAAAGAACTCAAGACTTCTGGGAAGTTTCAGTTAGGAATACCAC 2842
 QY 959 ACCCGCGCGGCTGAAGAAGAGAGCGCTGACCGTCTGACGTGGCGGCGACGCTACT 1018
 Db 2843 ACCCGCGAGGTTTAAAAAAGAAAAATCAGTAACAGTATTGGATGTGGTGTGATGCTACT 2902
 QY 1019 TCAGCGTCCCGCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCTCCCGCAGCATCA 1078
 Db 2903 TTTTCAGTTCCCTTAGATAAAGACTTTAGAAAGTATATCTGCAATTTACCATACCTAGTATA 2962
 QY 1079 ACAACGAGACCCCGGCACTCCGCTACCAAGTACAAGCTGCTGCCCCCGAGGCTGGAAAGGCA 1138
 Db 2963 ACAATGAGACACAGGAGTATAGATATACATAGTACAATGTCTGCCACAGGATGGAAAGAT 3022
 QY 1139 GCCCCAGCATTTCCAGAGCAGATGACCAAGATCTCGGAGCCCTTCCGCGCCCGCAACC 1198
 Db 3023 CACCGAGCAATATTCCAAAGTAGCATGACAAAAATCTTTAGAGCCTTTTGAAGAAACAGATC 3082
 QY 1199 CCGAGATGCTGATCTACCA- - - - -GGCCCCCTCTAGTGGCGGACGACCTGGAGATCG 1252
 Db 3083 CAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAATAAG 3142
 QY 1253 GCCAGCAGCGCGCAAGATCGAGGAGCTGCGCAAGCACTGCTGGCTGGGCTTTCACCA 1312
 Db 3143 GGCAGCATAGAACAAATAAGAGAACTGACAGCAGCATCTGTGAGGTGGGATTTACCA 3202
 QY 1313 CCCCCAGACAGAAGCAGAGGAGGCCCCCTTCTCTGATGGGCTACGAGCTGCAC 1372
 Db 3203 CACCAGACAAAAACATCAGAAAGAACCTCCATTCTTTGGATGGGTTATGAACCTCCATC 3262

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:45:48 ; Search time 107.667 Seconds
(without alignments)
10843.369 Million cell updates/sec

Title: US-09-610-313-31
Perfect score: 2463
Sequence: 1 gtcgacccaccatggccga.....gggctagcacccggtgaattc 2463
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues
Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PublishedApplications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US50_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US50_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1632.6	66.3	4307	10	US-09-999-183-2
2	1566	63.6	8366	10	US-09-872-733-6
3	1530.6	62.1	4338	10	US-09-872-733-1
4	1163.8	47.3	2467	10	US-09-872-733-3
5	1106.8	44.9	2601	10	US-09-735-487-7
6	1106.8	44.9	2601	10	US-09-735-487-9
7	1106.8	44.9	2601	10	US-09-735-487-11
8	1106.8	44.9	2601	10	US-09-735-487-13
9	1106.8	44.9	4307	10	US-09-999-183-1
10	1106.8	44.9	9719	10	US-09-737-190A-1
11	1054.4	42.8	8933	10	US-09-943-286-3
12	1054.4	42.8	8933	10	US-09-943-286-4
13	1054.4	42.8	8933	10	US-09-943-286-9
14	1048	42.5	9544	10	US-09-798-675-4
15	1048	42.5	9918	10	US-09-798-675-5
16	843.6	34.3	2348	9	US-10-097-997-1
17	831	33.7	9793	9	US-09-886-156-56
18	831	33.7	9793	9	US-09-886-150-56
19	831	33.7	9793	9	US-09-886-149-56

Sequence 56, Appl
Sequence 2, Appli
Sequence 14, Appl
Sequence 15, Appl
Sequence 5, Appli
Sequence 58, Appl
Sequence 24, Appl
Sequence 1, Appl
Sequence 4, Appli
Sequence 56, Appl
Sequence 1, Appli
Sequence 4, Appli
Sequence 1, Appli
Sequence 25, Appl
Sequence 16, Appl
Sequence 15, Appl
Sequence 1, Appli
Sequence 1, Appli
Sequence 17, Appl
Sequence 7, Appli
Sequence 15, Appl
Sequence 16, Appl
Sequence 17, Appl
Sequence 21, Appl
Sequence 13, Appl

831 33.7 9793 9 US-09-886-159-56
20 29.4 2507 10 US-09-872-733-2
21 682.4 27.7 12379 10 US-09-991-258-14
22 679.2 27.6 1323 10 US-09-991-258-15
23 450.6 18.3 948 10 US-09-756-551A-5
24 243.8 9.9 1800 9 US-10-003-035-58
25 243.8 9.9 1800 9 US-09-968-355-24
26 239.2 9.7 4516 10 US-09-968-355-24
27 216.4 8.8 12523 10 US-09-991-258-1
28 215.6 8.8 1476 10 US-09-991-258-4
29 172.6 7.0 507 9 US-10-003-035-56
30 165 6.7 9464 10 US-09-738-847-1
31 163.2 6.6 1482 10 US-09-818-443-4
32 163.2 6.6 1532 10 US-09-818-443-1
33 157.8 6.4 1503 10 US-09-968-355-25
34 156.6 6.4 1752 10 US-09-968-355-16
35 156.6 6.4 4594 10 US-09-968-355-15
36 155.8 6.3 13049 9 US-09-996-073-1
37 155.8 6.3 13049 10 US-09-231-235-1
38 155.8 6.3 13049 10 US-09-797-518A-1
39 133.6 5.4 1496 9 US-10-003-035-17
40 129.2 5.2 9751 10 US-09-946-239-7
41 118.2 4.8 2523 10 US-09-476-242-15
42 117.8 4.8 2517 10 US-09-476-242-16
43 117.8 4.8 2517 10 US-09-476-242-17
44 115.8 4.7 2310 10 US-09-476-242-21
45 115.8 4.7 2535 10 US-09-476-242-13

ALIGNMENTS

RESULT 1

US-09-999-183-2
; Sequence 2, Application US/09999183
; Patent No. US20020147169A1
; GENERAL INFORMATION:
; APPLICANT: MITROPHANOUS, et al
; TITLE OF INVENTION: In Vivo Selection Method
; FILE REFERENCE: 674523-2009
; CURRENT APPLICATION NUMBER: US/09/999,183
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: PCT/GB00/02136
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 9912965.2
; PRIOR FILING DATE: 1999-06-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: SeqWin99
; SEQ ID NO 2
; LENGTH: 4307
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Codon optimised gagpol sequence
US-09-999-183-2

Query Match 66.3%; Score 1632.6; DB 10; Length 4307;
Best Local Similarity 82.0%; Pred. No. 4.8e-263;
Matches 1947; Conservative 0; Mismatches 404; Indels 22; Gaps 5;
QY 12 CATGCCGAGGCCATGAGCCAGG---CCACCAGCGCCCAACATCTCTGATGACGCGCAGCAA 68
Db 1086 CTTGCTGAGGCCATGAGCCAGGTGACCAACTCCGCTACCATCATGATGACGCGCGCAA 1145
QY 69 CTTCAAGGCCCCCAAGCCCATCATCAAGTCTTCAACTGGCGCAGGAGGCCACATCGC 128
Db 1146 CTTTCGGAACCAACGCAAGATCGTCAAGTCTTCAACTGTGGCAAGAGGCGCACACAGC 1205
QY 129 CCGCAACTGCCCGCCCGCCCGCCCGAGGCTGCTGGAAGTGGCGCAAGGAGGCCACCA 188
Db 1206 CCGCAACTGCGAGGCCCGCTAGAAAGAGGCTGTGGAATGCGGCAAGAGGCCACCA 1265
QY 189 GATGAAGGATGTCACCGAGCGCCCGCAACTTCTCCGCGAGGACCTTGGCTTCCCCCA 248

Db 1266 GATGAAGACTGTACTAGAGACAGGCTAA-TTTTTTAGGAAGATCTGGCCCTTCTCTACA 1324
QY 249 GGGCAAGGCCGGGAGTTCCCGAGCAGAGAACCGCGCCAAACAGCCCGACAGCCCGCGA 308
Db 1325 AGGAAGGCCAGGGAATTTCTTCAGAGCAGAGCAGAGCCAAACAGCCCGACAGAGAGA 1384
QY 309 GCTGCAAGTGGCGGG-----CGACAAACCCCGCAGCGAGCGCGCGCGAGCGCCAGGG 362
Db 1385 GCTTCAGTCTGGGTAGACACAACACTCCCTCAGAAGCAGGAGCCGATAGACNAGG 1444
QY 363 CA-----CCCTGAACCTCCCGAGATCACCCCTGTGCGAGCGCCCGCTGTGAGCATCAA 416
Db 1445 AACTGTATCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAATAAA 1504
QY 417 GTTGGGGCGGCAGATCAAGAGAGCCCTGCTGGACACCGCGCGCGACACACCCGTGCTGGA 476
Db 1505 GATAGGGGGCAGCTCAAGAGAGCTCTCTCGACACCGGACAGACACCCGTGCTGGA 1564
QY 477 GGAGATGAGCTGCGCGCAAGTGGAAAGCCCAAGATGATCGGGGCAATCGCGCGCTTCAT 536
Db 1565 GGAGATGCTGTTGCCAGCCGCTGGAGCCCGAGATGATCGGGGATCGCGGTTTCAT 1624
QY 537 CAAGGTGGCCAGTACGACAGATCTCTGATCGAGATCTGCGGCAAGAGCCATFCGCGAC 596
Db 1625 CAAGGTGGCCAGTATGACAGATCTCTCATCGAAATCTGCGGCCACAAAGCTATCGGTAC 1684
QY 597 CGTCTGATCGGCCACCCCGTGAACATCATCGGCCCGCAACATGCTGACCCAGCTGGG 656
Db 1685 CGTCTGGTGGGCCACACCCGTCACATCATCGGACGAACTGTTGAGCGAGATCGG 1744
QY 657 CTGACCCCTGAACCTTCCCATCAGCCCATCGAGACCGTCCCGTGAAGCTGAAGCCCGG 716
Db 1745 TTGCACGCTGAACCTTCCCATTAGCCCTATCGAGACGTTACCGGTGAAGCTGAAGCCCGG 1804
QY 717 CATGGAGCCCGCAGTGAAGAGTGGCCCTGACCGAGGAGNAGATCAAGGCCCTGAC 776
Db 1805 GATGGAGCCCGAAGTCAAGCAATGGCCATTGACAGAGAGAGAGATCAAGGCACCTGGT 1864
QY 777 CGCCATCTCGGAGAGATGAGAGAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCC 836
Db 1865 GGAGTTTGACAGAGATGGAAGAGGAGGAAATCTCCAAGATTGGGCTGAGAACCC 1924
QY 837 CTACAACCCCGCTGTTGCGCATCAAGAGAGAGGACAGCACCAAGTGGGCGCAAGCTGGT 896
Db 1925 GTACAACACGCGGTGTTGCGCAATCAAGAGAGAGGACTCGACGAAATGGGCGCAAGCTGGT 1984
QY 897 GGACTTCGCGAGTGAACAGCGCACCCAGGACTTCTGGAGGTGAGCTGGGCACTCC 956
Db 1985 GGACTTCGCGAGTGAACAGCGCACCGCAAGACTTCTGGAGGTTGAGCTGGGCACTCC 2044
QY 957 CCACCCCGCGGCTGAAGAGAGAGAGCGTGACCGTGTGACGCTGGGCGACGCTA 1016
Db 2045 GCACCCCGAGGCTGAAGAGAGAAATCCGTGACCGTACTGGATGGGTGATGCTA 2104
QY 1017 CTTGAGGCTGCGCTGGACGAGGACTTCCGCAAGTACACCGCCCTTACCATTCGCCAGAT 1076
Db 2105 CTTCTCGCTTCCCTGGACGAAGACTTCAGGAAGTACACTGCTTCAATTCCTCGAT 2164
QY 1077 CAACAAGAGACCCCGCATCCGCTACCACTACAAGTGTGTCGCCCGCAGGCTGGAAGG 1136
Db 2165 CAACAAGAGACCCGGGATTCGATATGATACAGAGTGTGTCGCCCGCAGGCTGGAAGG 2224
QY 1137 CAGCCCGCAGCATCTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCCGCCCGCAA 1196
Db 2225 CTCTCCGCAATCTCCAGAGTAGCATGACCAAAATCCTGGAGCCTTTCGCGCAACAGAA 2284
QY 1197 CCGCGAGATCGTATCTACCA-----GGCCCCCTGTACGTGGGACCGACCTGGAGAT 1250
Db 2285 CCGCGACATCGTATCTATCATAGTACATGATGACTTGTACGTGGGCTGTGATCTAGAGAT 2344
QY 1251 CGGCGACCGCGCCCAAGATCGAGGAGCTGCGCAAGACCTGTGCTGCGCTGGCTTCAC 1310
Db 2345 AGGGACGACCGCCACCAAGATCGAGGAGCTGCGCCAGCACCTGTGTAGGTGGGACTGAC 2404

QY 1311 CACCCCGCACAAAGCACCAAGAGAGCCCCCTTCTCTGTGGATGGCTTACGAGCTGCA 1370
Db 2405 CACACCCGACAAAGACCAAGAGAGCCCTTCTCTGTGGATGGCTTACGAGCTGCA 2464
QY 1371 CCCGACAAGTGGACCGTGCAGCCCATCGAGTGCCTCGAGAAAGAGAGAGCTGGACCGTGA 1430
Db 2465 CCCTGACAAATGGACCGTGCAGCTTATCGTGTGCCAGAGAAAGACAGCTGGAGTGTCAA 2524
QY 1431 CGACATCCACAAGCTGTTGGGCAAGCTGAATGGCCAGCCAGATCTACCCCGCATCAA 1490
Db 2525 CGACATACAGAAGCTGTTGGGGAAGTTGAATGGGCCAGTCAGATTTACCCAGGGATTA 2584
QY 1491 GGTGCGCACCTGTGTGAAGCTGTCGCGCGCCCAAGGCCCTGACCCACATCTGTCGCCCT 1550
Db 2585 GTTGAGCAGCTGTGCAAACTCTCCCGGACCAAGGCACTACAGAGTGTATCCCCCT 2644
QY 1551 GACCGAGAGCGGAGCTGGAGTGGCCGAGAACCCGAGATCTGTCGCGAGCCCGTGA 1610
Db 2645 AACCGAGAGCGGAGCTCGAACTGCGAGAAACCGAGAGATCTTAAGAGGAGCCGTGA 2704
QY 1611 CGCGGTGTACTAGACCCAGCAGAGGACCTGGTGGCCGAGATCCAGAGCAGGCGCCACGA 1670
Db 2705 CGCGGTGTACTAGACCCCTCCAAAGGACCTGATCGCCGAGATCCAGAGCAGGCGCAGG 2764
QY 1671 CCAGTGGACCTTACCAGATCTACCAGAGGACCTTCAAGAACCTGAAGACCGGCAAGTACGC 1730
Db 2765 CCAGTGGACCTATCAGATTTACCAGAGGACCTTCAAGAACCTGAAGACCGGCAAGTACGC 2824
QY 1731 CAGATCGGACCCCGCACACACAGAGCTGAAGACCTGACCCAGGCGCTGCGAGAGAT 1790
Db 2825 CCGGATGAGGCGTCCCGACACATAACAGCTCAAGCAGCTGACCCAGGCGCTGCGAGAGAT 2884
QY 1791 CGCCATGAGAGCATCTGTGATCTGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAA 1850
Db 2885 CACCACCAAGACATCTGTGATCTGGGAAAGACTCTTAAGTTCAAGCTGCCATCCAGAA 2944
QY 1851 GGAGACCTGGGAGACCTGTTGGACCGACTACTGCGAGGCCACTTGGATTCGCCGAGTGGGA 1910
Db 2945 GGAAACCTGGGAAACCTGTTGGACAGAGATTGGCAGGCCACCTGGATTCTCGAGTGGGA 3004
QY 1911 GTTCGTGAACACCCCGCCCTTGGTGAAGCTGTGTACCAGCTGGAGAGGAGCCCATCAT 1970
Db 3005 GTTCGTGAACACCCCTTCCCTGGTGAAGCTGTGTACCAGCTGGAGAGGAGCCCATAGT 3064
QY 1971 CGGCGCGAGACCTTCTACGTGGAGCGCGCCCAACCGCAGACCAAGATTCGGCAAGGC 2030
Db 3065 GGGCGCGAGACCTTCTACGTGGATGGGCGCTTAACAGGAGACTAAGCTGGGCAAGC 3124
QY 2031 CGGCTACGTGACCAACCGGCGCGCAGAAATCTGTAGCCTGACCGAGACCAACCA 2090
Db 3125 CGGATACGTCACTAACCGGCGCAGACAGAGTGTTCACCCCTCACTGACACCAACCA 3184
QY 2091 GAACACCGAGCTGACGCCATCCAGCTGGCCCTGCAGACAGCGCAGAGTGAACAT 2150
Db 3185 GAAGCTGAGCTGACGCCATTTTACCTGCTTTCGAGGACTCGGGCTTGGAGGTGAACAT 3244
QY 2151 CGTACCGACAGCAGCAGTACGCCCTGGGCATCATCCAGGCCCGCCGACAGAGCGAGAG 2210
Db 3245 GTCGACAGCTCTCAGTATGCCCTGGGCATCATTAAGCCCGCAGCAGACAGAGTGTG 3304
QY 2211 CGAGCTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGAGTGTACCTGAGCTG 2270
Db 3305 CGAGCTGTGTAATCAGATCATCGAGCAGCTGATCAAGAGGAGGAGTCTATCTGCGCTG 3364
QY 2271 GGTCCCGCCCAAGGCGCATCGGCGCAACGAGCAGATCGACAGCTGGTGGAGCAAGG 2330
Db 3365 GGTACCCCGCCCAAGGCGCATTTGGGCGCATGAGCAGCTGCAACAGCTGCTCGGCTGG 3424
QY 2331 CATCCGCAAGTGTGTTCTTCCTGGACCGCATCGA 2363
Db 3425 CATCAGGAAGTGTATTTCCTGGATGGCATCGA 3457

RESULT 2

US-09-872-733-6
; Sequence 6, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; FILE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCI/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of the construct pcmvagagpolBKN containing a CMV
; OTHER INFORMATION: promoter, a HIV gag/pol gene and a kanamycin
; OTHER INFORMATION: resistance gene
US-09-872-733-6

Query Match 63.6%; Score 1566; DB 10; Length 8366;
Best Local Similarity 80.3%; Pred. No. 5.6e-252;
Matches 1905; Conservative 0; Mismatches 445; Indels 22; Gaps 5;
QY 14 TGGCCGAGGCGCATGAGCAGGCCACCAGC---GCCAATCTCTGATGAGCCGACCACT 70
DB 1857 TGGCCGAGGCGATGAGCCAGGTGCGAATCTGGCGACCAATATGATGAGAGGCGAAT 1916
QY 71 TCAAGGCCCCAAGCGCATCAATCAAGTCTTCACTGGGGCAAGGAGGCCACATCGCCC 130
DB 1917 TCGGAACACAGCGGAAGATCGTCAAGTGTCTCAATGTGGCAAGAGGCGACACGCCA 1976
QY 131 GCAACTGCGCGCCCGCCCGCAGAGAGGCTGCTGGAAGTGGCGGCAAGGAGGCCACAG 190
DB 1977 GAACTTGGCGGGCCCCCGGGAAGAGGGCTGTGGAATGTGGAAGAGGAGCACACAAA 2036
QY 191 TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTCCGCGAGGAACTGCGCTTCCCGCAGG 250
DB 2037 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTATGGAAGATCTGGCTTCTTACAG 2095
QY 251 GCAAGCGCGGAGTTCCTCCAGCGAGCAGAAACCGGCCAACAGCCCGCCAGCCCGAGC 310
DB 2096 GGAAGGCCAGGGAATTTTCTTCAGAGCAGACCCAGAGCCAAACAGCCCGCCAGAGAGC 2155
QY 311 TCCAGGTGCGGG-----CGACACCCCGCAGCGAGCGCGCGCGCGCGCGCGCGCGGCA 364
DB 2156 TTCAGGTCTGGGGTAGAGACAACTCTCCCTCAGAGCAGAGGAGCGCGATAGACAGGAA 2215
QY 365 -----CCTGAACTTCCCGCAGATCACTCTGTGCGAGCGCCCTGTGTGAGCATCAAGG 418
DB 2216 CTGTATCTTTAACTTCCCTCAGATCATCTTTTGGCAACAGCCCTCGTCAGATGAAGA 2275
QY 419 TGGCGCGCCAGATCAAGGAGGCCCTTCTGGACACCGCGCGCGCGCGCGCGCGCGCGGAG 478
DB 2276 TCGGGGGCAACTCAAGGAAGCGTGTCTGATACAGAGCAGATGATACAGTATTAGAG 2335
QY 479 AGATGAGCCTGCCCGCAAGTGGAGGCCAAGATGATCGGGGCGATCGCGGCGCTTCATCA 538
DB 2336 AAATGAGTTTCCAGGAAGATGGAACCAAAATGATAGGGGGGATCGGGGCTTCATCA 2395
QY 539 AGGTGCGCGAGTACGACCATCTCTGATCGAGATCTCGCGCAAGAGGCCATCGGCACCG 598
DB 2396 AGGTGAGCGAGTACGACCATCTCTATAGAAATCTGTGGACATAAAGCTATAGGTACAG 2455

QY 599 TGCTGATCGGGCCCCAGCCCGCTGAACATCATCGCGCRAACATCTGCTGACCCAGCTGGCT 658
DB 2456 TATTAGTAGGACCTACACCTGTCAACATATTTGAGAGAAATCTGTTGACCCAGATCGGCT 2515
QY 659 GCACCTCTGAATTTCCCATCATGAGCCCATCGAGACCCTGCGCGTGAAGCTGAAGCCCGGCA 718
DB 2516 GCACCTTGAATTTCCCATCATGAGCCCTATTGAGACGGTGCCTGAACTTGAAGCCGGGA 2575
QY 719 TGGACCGCCCAAGGTGAGCAGTGGCCCTGACCCAGAGAGAGATCAAGGCCCTGACCG 778
DB 2576 TGGACCGCCCAAGGTGAGCAGTGGCCCTGACCAATGACAAAGAGAGATCAAGGCCCTTAGTC 2635
QY 779 CCATCTCCGAGGAGATGGAGAGGAGGCAAGATCACCAAGATCGCGCCCGGAGAACCCCT 838
DB 2636 AAATCTGTACAGAGATGGAGAGGAGGAGAGATCAGCAGATCGGCGCTGAGAACCCCT 2695
QY 839 ACAACACCCCGTGTTCGCATCAAGAAGAAGACAGCACCAGTGGCGCAAGCTGGTGG 898
DB 2696 ACAACACTCCAGTCTTCGCAATCAAGAAGAAGACAGTACCAAGTGGAGAAAGCTGGTG 2755
QY 899 ACTTCCGCGAGCTGAACAAGCGCCAGGACTTCTGGAGGTGCGAGTGGGATCCCC 958
DB 2756 ACTTCAGAGAGCTGAACAAGAACTCAGGACTTCTGGGAAGTTTCACTGGGATCCCC 2815
QY 959 ACCCCCGCGCTGAAGAAGAAGAGAGCTGACCTGTGAGAGTGGCGAGCGCCCTACT 1018
DB 2816 ATCCCGCTGGTTGAAGAAGAAGTCACTGATGACAGTCTGGATGTGGGTGATGCTACT 2875
QY 1019 TCAGGCTGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACATCCCCAGCATCA 1078
DB 2876 TCTCCGTTCCCTTGGACGAGGACTTTCAGGAAGTACACTGCCCTCAGCATACCTAGCATCA 2935
QY 1079 ACAACGAGACCCCGCGCATCCGCTACAGTACACAGTGTGTGCCCGGAGGCTGGAAGGCA 1138
DB 2936 ACAACGAGACCCCGCGCATCCGCTACAGTACACAGTGTGTGCCCGGAGGATGGAAGGAT 2995
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATCACCAAGATCTGAGAGCTTCCGCGCCGCAAC 1198
DB 2996 CACGAGCATCTTCAAGCAGCATGACCAAGATCTGTGAGGCTTCCGCAAGCAAAAC 3055
QY 1199 CGAGATCTGTATCTACCA-----GGCCCCCTGTGCTGGGCGAGCAGCTGGAGATCG 1252
DB 3056 CAGACATCTGTATCTATCATGATACATGGAGCACTCTACCTAGGAAGTACCTGGAGATCG 3115
QY 1253 GCGACGCGCGCGCAGATCGAGGAGCTGCGAAGACACCTGTGCGCTGGGCTTTCACCA 1312
DB 3116 GCGACGACAGGACCAAGATCGAGGAGCTGAGACAGCATCTGTGAGGTGGGACTGACCA 3175
QY 1313 CCGCGGCAAGAGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1372
DB 3176 CACGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3235
QY 1373 CCGACAAGTGGACGCTGAGCCCATGAGCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1432
DB 3236 CTGACAAGTGGACGCTGAGCCCATGAGCTGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAG 3295
QY 1433 ACATCCAGAGCTGTGGCGAGCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1492
DB 3296 ACATACAGAGCTGTGGGCAAGTGAATGGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3355
QY 1493 TCGCGCAGCTGTGCAAGCTGCTGCGCGGGGCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1552
DB 3356 TTAGGCGAGCTGTGCAAGCTGCTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3415
QY 1553 CCGAGAGGCGAGCTGAGCTGGCGGAGAACCGGAGATCTCTGCGCGAGGAGGAGGAGGAGGAGGAG 1612
DB 3416 CAGAGGAAGCAGAGCTAGAACTGGCAGAGAGACCGAGAGATCTTGAAGGAGGAGGAGGAGGAG 3475
QY 1613 GCGTGTACTACGACCCCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1672
DB 3476 GAGTGTACTACGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3535
QY 1673 AGTGGACCTACGAGATCTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1732

Db 3536 AATGGACCTACCAAAATCTACCGAGGCGCTTCAAGAACCTTGAAGACAGGCAAGTACGCAA 3595
QY 1733 AGATGGCACCAGCCACACCAACGAGCTGAAGCAGCTGACCGAGGCGCTGCAGAAATCG 1792
Db 3596 GGATGAGGGGTGCCACACCAACGATGTGAAGCAGCTGACAGAGGCGAGTCAAGAATCA 3655
QY 1793 CCATGAGAGCATCTGATCTGGGGGAAGACCCCAAGTTTCCGCTGCCCATCCAGAGG 1852
Db 3656 CCACAGAGACATCTGATCTGGGGCAGAGCTCCCAAGTTTCAAGTTTCAAGCTGCCATACAGAGG 3715
QY 1853 AGACCTGGGAGACATCTGGTGAACGACTTACTTGGCAGGCCACCTTGGATCCCGCAGTGGAGT 1912
Db 3716 AGACATGGGAGACATCTGGTGAACGAGTACTTGGCAAGCCACCTGGATCCCTGAGTGGAGT 3775
QY 1913 TCGTGAACACCCCGCTTGGTGAAGCTGTGGTACCAAGCTGGGAGAGGCCATCATCG 1972
Db 3776 TCGTGAACACCCCGCTTGGTGAAGCTGTGGTATCATCAAGCTGGGAGAGGCCATCATCG 3835
QY 1973 GCGCGGAGACCTTCTAGCTGGAGCGCGCCCAACCGCAGACCAAGATCGGAAGCGCG 2032
Db 3836 GAGCAGAGACCTTCTAGCTGGATGGGCGAGCCACACAGGGAGACCAAGCTGGGCAAGCGAG 3895
QY 2033 GCTAGCTGACCGACCGGGCGGCGGAGAGATCGTGAAGCTTGAAGCAGACCAACCAACGAGA 2092
Db 3896 GCTAGCTGACCAACCGAGGAGCAGACAGAAAGTGTGAAGCTTGAAGCTTGAAGCTTGAAGCT 3955
QY 2093 AGACCGAGCTGACGAGGACCTTACCTGAGCTGAGGAGCAGCGAGGAGTGAACATCG 2152
Db 3956 AGCTGAGCTGCAAGCCATCTACCTAGCTTGAAGAGCAGCGAGTGAAGTGAACATCG 4015
QY 2153 TGACCGAGCAGCAGTACGCTTGGGCTATCTCCAGGCGCCAGCAGCAGCAGCAGCAGCAGC 2212
Db 4016 TGACAGACTCACAGTACGCTTGGGCTATCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4075
QY 2213 AGCTGGTGAACAGATCATCGACAGCTGATCAAGAGAGGAGAGTGTACTGAGCTGGG 2272
Db 4076 AGCTGGTGAACAGATCATCGACAGCTGATCAAGAGAGGAGAGTGTACTGAGCTGGG 4135
QY 2273 TGCCCGCCCAACAGGAGCTGCGGCGCAACGAGCAGATCGCAAGCTGGTGAAGAGGCA 2332
Db 4136 TACAGCAGCAGCAAGAAATGGAGAAATGACAAAGTAGATAAATTAGTCACTGCTGGGA 4195
QY 2333 TCCGCAAGGTGCTTCTTCCAGGCGCATCGAT 2364
Db 4196 TCCGCAAGGTGCTTCTTCCAGGCGCATCGAT 4227

RESULT 3

US-09-872-733-1
; Sequence 1, Application US/09872733
; Patent No. US2001003665A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4338
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Gag/Pol gene
US-09-872-733-1

Query Match 62.1%; Score 1530.6; DB 10; Length 4338;
Best Local Similarity 80.1%; Pred. No. 4e-246;
Matches 1905; Conservative 0; Mismatches 444; Indels 28; Gaps 8;

QY 14 TGGCCGAGGCGCATGAGCCAGGCGCACCAAGC---GCCAACATCTCTGATCGAGGCGAGCAACT 70
Db 1085 TGGCCGAGGCGCATGAGCCAGGCGCACCAAGC---GCCAACATCTCTGATCGAGGCGAGCAACT 1144
QY 71 TCAGGGCCCCCAAGCGCATCATCAAGTGCCTCAACTCGGCAAGGAGGCGCACATCGGCC 130
Db 1145 TCCGGAACCAAGCGAGATCGTCAAGTGCCTCAATTTGGCAAGAGAGGCGCACACCGCCA 1204
QY 131 GCAACTCCCGCGCGCCCGCCCAAGAAAGGCGCTCTGGAAGTGGCGCAAGGAGGCGCACCAAG 190
Db 1205 GGAAGTCCCGCGCGCCCGCCCAAGAAAGGCGCTCTGGAAGTGGCGCAAGGAGGAGCACCAAA 1264
QY 191 TGAAGAGCTACAGCGAGCGCCAGGCGCAACTTCTCCGCGAGGAGCTGGCTTCCCGCAGG 250
Db 1265 TGAAGAGTCTTACTTGAGAGAGAGGCTAA-TTTTTTAGGAGAGATCTGGCCTTCTCAAG 1323
QY 251 GCAGGCGCGCGAGTTCGCCAGCGAGCAGACCGCGCAACAGCGCCACAGCGCCAGCGCAGC 310
Db 1324 GGAAGCGCGAGGAAATTTCTTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1393
QY 311 TGCAGGTGCGCGG-----CGACAACCGCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCA 364
Db 1384 TGCAGGTGCGCGGTTAGAGCAACAACCTCCCGCTCAGAAGCAGGAGCGCATAGACAAGGAA 1443
QY 365 -----CCCTGAACTTCCCGCAGATCACTCTTGGCAACAGCAGCGCTCTGACAGTAAGGA 418
Db 1444 CTGTATCTTTAACTTCCCTCAGATCACTCTTGGCAACAGCAGCGCTCTGACAGTAAGGA 1503
QY 419 TGGCGCGCGCAGATCAAGAGGCGCTCTGTCAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 478
Db 1504 TCGGGGGGCAACTCAAGAGAGCGCTGCTGATACAGAGCAGATGATACATATTAGAG 1563
QY 479 AGATGAGCCTTCCCGCGCAAGTGGAAAGCCAAAGATGATCGCGCGCGCGCGCGCGCGCGCGCG 538
Db 1564 AATGAGTTCGCGAGGAGATGGAAACCAAAATGATAGGGGGGATCGGGGCGCTTCATCA 1623
QY 539 AGTTCGCCAGTACGACAGATCTGATCGAGATCTCGGCAAGAGGCGCGCGCGCGCGCGCGCG 598
Db 1624 AGTTCGCCAGTACGACAGATCTGATCGAGATCTGTCAGACATTAAGCTATAGGTACAG 1683
QY 599 TGCTGATCG 654
Db 1684 TATTAGTAGGACCTTACCTACACCTGTCAACATATTGGAAGAAATCTGTTGACCGCAGATC 1743
QY 655 GGCTGACCGCTGAACTTCCCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 714
Db 1744 GGCTGACCGCTTGAACCTTCCCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1803
QY 715 GGCATGACG 774
Db 1804 GGCATGACG 1863
QY 775 ACCGCCATCTCGGAGGAGATGGAGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 834
Db 1864 GTCGAATCTGTACAGAGATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1923
QY 835 CCCTTACAAACCG 894
Db 1924 CCCTTACAAACCG 1983
QY 895 GTGGACTTCCCG 954
Db 1984 GTGGACTTCCCG 2043
QY 955 CCCCACCG 1014
Db 2044 CCACATCCCG 2103

QY 1015 TACTTCAGCGTCCCGTGGAGCGAGGACTTCGCAAGTACACCGCTTTCACCATCCCGAGC 1074
Db 2104 TACTTCCGTTCCCTTGGAGGAGACTTCAGGAAGTACACTGCTTTCAGGATACCTAGC 2163
QY 1075 ATCAACAACGAGAGCCCGCGATCCCGTACAGTACAACTGCTGCGCCCGAGGGGTGAAG 1134
Db 2164 ATCAACAACGAGAGACACGAGGATCCCGTACAGTACAACTGCTGCGCCCGAGGGATGAAG 2223
QY 1135 GGCAGCCCGAGCATCTTCCAGAGCAGCATGACCAAGATCTGAGAGCCCTTCCGCGCCCGC 1194
Db 2224 GGATCACCGACCATCTTCAAGCAGCATGACCAAGATCTGAGAGCCCTTCCGCAACGAA 2283
QY 1195 AACCCCGAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGGACGACCTGGAG 1248
Db 2284 AACCCAGACATCGTGATCTATCAGTACATGAGCAACCTCTAGGTAGGAAGTGAACCTGGAG 2343
QY 1249 ATC-GGCCAGCAGCGCCCGAGATCGAGGAGCTGCGCAAGCACTGCTGCGCTGGGCGTT 1307
Db 2344 ATCGGGGCGACACAGGACCAAGATCGAGGAGCTGAGACAGCATCTGTTGAGGTGGGACT 2403
QY 1308 CACCACCCCGACAAAGACACCAAGAGGAGCCCGCTTCTGTGGATGGGCTACGAGCT 1367
Db 2404 GACCACACGAGCAAGAGACCAAGAGGAGGACCTTCTGTGGATGGGCTACGAGCT 2463
QY 1368 GCACCCCGACAAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGGAGAGCTGGACCGT 1427
Db 2464 GCATCTTGACAAGTGGAGACGTGCGAGCCCATCGTCTGCTGAGAGGAGACAGCTGAGCTGT 2523
QY 1428 GAAGGACATCCAGAGCTGGTGGCAAGCTGAAGTGGGCGAGCAGATCTACCCCGGAT 1487
Db 2524 GAAGGACATACAGAGCTGGTGGCAAGTGAAGTGGGCGAGCAGATCTACCCAGGAT 2583
QY 1488 CAAGGTGGCGAGCTGTGAAGTGTGCGCGCGCAAGGCGCTGACCGACATCTGTGCC 1547
Db 2584 CAAGTTAGGAGCTGTGAAGTGTGCGCGCGCAAGGCGCTGACAGAGTGTATCCC 2643
QY 1548 CCTGACCGAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTTGGCGGAGCCCGT 1607
Db 2644 ACTGACAGAGGAGCAGAGCTAGAACTGGCAGAGAACCGGAGATCTGTAAGGAGCCAGT 2703
QY 1608 GCAGGCGGTACTACAGCCCGCAGCAGGAGCTGCTGGCGGAGATCCAGAGCAGGCGCA 1667
Db 2704 ACATGGAGTGTACTACAGCCCGCAGCAGGAGCTGATCGCAGAGATCCAGAGCAGGCGCA 2763
QY 1668 CGACCATGGACCTTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTA 1727
Db 2764 AGGCAATGAGCTTACCAATCTACAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTA 2823
QY 1728 CGCAAGATCGCACCGCCCGCAGCAGCTGAAGCAGCTGACCGAGGCGCTGCGAGAA 1787
Db 2824 CGCAAGATGAGGCGTCCCGCAGCAGCAGCTGTAAGCAGCTGACAGGCGAGTGCAGAA 2883
QY 1788 GATCGCATGGAGAGCATCTGTGATCTGGGCAAGACCCCGAAGTTCGCGCTGCCCTTCCA 1847
Db 2884 GATCACCACAGAGAGCATCTGTGATCTGGGCAAGACTCCCAAGTTCAGAGCTGCCATACA 2943
QY 1848 GAAGGAGACCTTGGAGACCTTGGTGGACCGCTACTGCGAGCCAGCTGGATCCCGGAGTG 1907
Db 2944 GAAGGAGACATGGAGAGCATGTTGGACCGAGTACTGCAAGCCAGCTGGATCCCTGAGTG 3003
QY 1908 GGAGTTCTGTAACACCCCGCTGGTGAAGCTGTGTGATGCTGAGCTGGAGAGGAGCCCAT 1967
Db 3004 GGAGTTCTGTAACACCCCTTCCCTTGGTGAAGCTGTGTGATGCTGAGCTGGAGAGGAGCCAT 3063
QY 1968 CATCGGCGCGAGACCTTCTACGTGGAGGCGCGCGCAACCGCAGAGCAAGATCGGCA 2027
Db 3064 CGTGGGAGCAGAGACCTTCTACGTGGATGGGCGAGCCAGCAAGGAGAGCCAGCTGGGCA 3123
QY 2028 GCGCGCTAGCTGACCGAGCGCGCGCAGAGATGCTGAGCTGACCGGAGACCCCA 2087
Db 3124 GCGAGGCTAGCTGACCGAGCGCGCGCAGAGAGTGGTGGACCTGACTGACACCCCA 3183
QY 2088 CCAGAGACCGAGCTGCGAGGCGCATCCAGCTGGCGCTGCGAGGAGCGGAGGTTGAA 2147

Db 3184 CCAGAAGACTGAGCTGCAAGCCATCTACCTAGCTCTGCAAGACAGCGGACTGGAAGTCAA 3243
QY 2148 CATCGTGACCGACAGCCAGTAGCGCCCTGGGCATCATCCAGGCCCGCCGACAGAGCGGA 2207
Db 3244 CATCGTGACAGACTTCAGAGTACG-CATGGGCATCATCCAAAGCAACAGACCAATCCGA 3302
QY 2208 GAGGAGCTGGTGAACAGAGCATCATCGAGCAGCTGATCAAGAGGAGGAGGTGTACCTGAG 2267
Db 3303 GTCAGAGCTGCTGAACAGAGCATCATCGAGCAGCTGATCAAGAGGAGGAGGTGTACCTGGC 3362
QY 2268 CTGGTCCCGCCCGACAAAGGCGATCGCGGCAACGAGCAGATCGACAAAGCTGGTGAGCAA 2327
Db 3363 ATGGGTACCAACACACAAAGGAATGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGC 3422
QY 2328 GGGCATCCGCAAGGTGCTGTTCTCTGGAGCGGATCGAT 2364
Db 3423 TGGATCCGGAAGGTGCTGTTCTCTGGAGCGGATCGAT 3459

RESULT 4
US-09-872-733-3
; Sequence 3, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Pol gene
US-09-872-733-3

Query Match 47.3%; Score 1163.8; DB 10; Length 2467;
Best Local Similarity 84.0%; Pred. No. 3e-185;
Matches 1328; Conservative 0; Mismatches 247; Indels 6; Gaps 1;
QY 790 GAGATGGAGAGGAGGCGCAAGATCACCAGAGATCGCCCGGAGAACCCCTACACACCCGCC 849
Db 7 GAGATGGAGAGGAGGAGGAGATCAGCAAGATCGGGCGCTGAGAACCCCTACACACTCCA 66
QY 850 GTGTTCGCATCAAGAAGAGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCGCGAG 909
Db 67 GTCTTCGCATCAAGAAGAGACAGTACCAAGTGGAGGAGGAGTGGTGGACTTCAGAGAG 126
QY 910 CTGAACAGCCAGCCAGGACTTCTGGAGGTGAGCTGGGCGATCCGCCAGCCCGCGGC 969
Db 127 CTGAACAGAGAACTCAGGACTTCTGGGAAGTTCAGCTGGGATCCCATCCCGCTGGG 186
QY 970 CTGAAGAGAGAGAGCGCTGACCGCTGAGCGTGGGCGAGCGCTACTTTCAGCGTGGCC 1029
Db 187 TTGAAGAGAGAGAGTCACTGACAGTGTGTGATGTGGTGTGATGCTTCTCGGTCCC 246
QY 1030 CTGACGAGGAGACTTCCGCAAGTACACCCCTTACCATCCCGAGCATCAACACAGAGACC 1089
Db 247 TTGACGAGGAGACTTCAGGAAGTACACTGCTTACGATACCTAGCATCAACACAGAGACA 306
QY 1090 CCCGGCATCCCGTACCAAGTACAAAGTGTGCTGCCCGAGGCTGGAAGGAGCGCCCGCATC 1149
Db 307 CCAGGATCCCGTACCAAGTACAAAGTGTGCTGCCACAGGATGGAAGGATGCAAGGATCACCAGCCATC 366

QY 1150 TTCAGAGCAGCATGACCAAGATCCTTGAGCCCTTCGCGGCCCCGCAACCCCGAGATCGTG 1209
Db 367 TTTCAAGCAGCATGACCAAGATCCTTGAGCCCTTCGCGAAGCAAAACCCAGACATCGTG 426
QY 1210 ATCTACCA-----GGCCCCCTTAGTGGGAGCAGCCTGGAGATCGGCCAGCACCGC 1263
Db 427 ATCTATCAGTACATGAGGAGCCTTACGTAGGAAGTACCTGGAGATCGGGCAGCAGG 486
QY 1264 GCCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGGCTTCACACCCCGACAAG 1323
Db 487 ACCAAGATCGAGGAGCTGAGACAGCATCTGTTGAGTGGGGACTGACCACACAGACAAG 546
QY 1324 AAGCACCAGAGAGCCCCCTTCCTGTGATGGGCTACGAGCTGCACCCCGACAAGTGG 1383
Db 547 AAGCACCAGAGAGAACCTCCCTTCCTGTGATGGGCTACGAACCTGCATCTCGACAAGTGG 606
QY 1384 ACCGTGAGCCCATCGAGCTGCCGAGAGAGAGAGCTGGAGCGTGAACAGATCCAGAAG 1443
Db 607 ACAGTGCAGCCCATCGTGTGCTGCGCTGGAAGGACAGCTGGAGCTGTGAACGACATACAGAAG 666
QY 1444 CTGTGGGCAAGCTGAATCGGGCGGCAAGCCCTGACGATCTACCCCGCATCAAGGTGCGCAGCTG 1503
Db 667 CTGTGGGCAAGTGAATCGGCAAGCCAGATCTACCCAGGCATCAAAAGTTAGCAGCTG 726
QY 1504 TGCAGCTGTGCGCGGCGCAAGGCCCTGACGACATCTGTCGCCCTGACCGAGAGGCC 1563
Db 727 TGCAGCTGTGCGCGGCGCAAGGCCCTGACGACATCTACCCAGGCATCAAAAGTTAGCAGCTG 786
QY 1564 GAGCTGAGCTGCGCGGAGAACCGGAGATCTGCGCGAGCCCTGACGCGGGTGTACTAC 1623
Db 787 GAGCTGAGCTGCGCGGAGAACCGGAGATCTGCGCGAGCCAGTACATGAGGTGTACTAC 846
QY 1624 GACCCAGCAGGAGCTGTGGCGGAGATCCAGAAGCAGGGCCAGCAGCTGACCTAC 1683
Db 847 GACCCAGCAGGAGCTGTGGCGGAGATCCAGAAGCAGGGCCAGCAGCTGACCTAC 906
QY 1684 CAGATCTACGAGGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATCGCGACC 1743
Db 907 CAAATCTACGAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATCGAGGGT 966
QY 1744 GCCACACCAAGCAGCTGAGCAGCTGACCGAGCGCTGACGAGATCGCCATGGAGAGC 1803
Db 967 GCCACACCAAGCAGCTGAGCAGCTGACCGAGCGCTGACGAGATCCACAGAGAGC 1026
QY 1804 ATCGTATCTGGGCGAAGACCCCAAGTTCCGCTGCGCCATCCAGAGGAGACCTGGGAG 1863
Db 1027 ATCGTATCTGGGCGAAGACTCCCAAGTTCAAGCTGCCATACAGAGGAGACATGGGAG 1086
QY 1864 ACCTGTGGACCGACTACTGGCAGGCGACCTGGATCCCGGAGTGGGAGTTCGTGAACACC 1923
Db 1087 ACATGTGGACCGAGTACTGGCAGGCGACCTGGATCCCGGAGTGGGAGTTCGTGAACACC 1146
QY 1924 CCCCCCTGTGTAAGCTGTGTTACGAGCTGAGAGAGCGCCATCATCGCGCGCGAGACC 1983
Db 1147 CTTCCCTGTGTAAGCTGTGTTACGAGCTGAGAGAGCGCCATCATCGTGGGAGCAGAGACC 1206
QY 1984 TTCTAGCTGACGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGGCTACGTGACC 2043
Db 1207 TTCTAGCTGATGGGCGAGCAACAGGAGACCAAGCTGGCAGGCGAGCGCTACGTGACC 1266
QY 2044 GACCGGGCGGCGAGAGATCTGTGAGGCTGACCGAGACACCAACAGAGACCGAGCTG 2103
Db 1267 AACCGAGGAGCAGAGAAGTGGTGAACCTGACCTGACCTGACACCAACCAAGAGACTGAGCTG 1326
QY 2104 CAGGCATCCAGCTGGCCCTGCGAGGAGCAGCGGAGCTGAACATCTGTCGACCGAGC 2163
Db 1327 CAGGCATCTACCTAGCTCTGCAAGACAGCGGAGCTGGAAGTGAACATCTGTCGAGACTCA 1386
QY 2164 CAGTACCGCTGGGCAATCAATCCAGGCGCCAGCCGCAAGAGCAGGAGCGAGCTGGTGAAC 2223
Db 1387 CAGTACGCACTGGGCAATCAATCCAGGAGCAGACCAATCCCGAGTCAAGAGCTGGTGAAC 1446

QY 2224 CAGATCATCGAGCAGCTGATCAAGAAGGAGAAAGTGTACTGTGAGCTGGTCCCGCCAC 2283
Db 1447 CAGATCATCGAGCAGCTGATCAAGAAGGAGAAAGTGTACTGTGATGGTACGACACAC 1506
QY 2284 AAGGGCATCGCGGCAAGCAGCAGATCGACAAGCTGGTGTGACAAGGGCATCCGCAAGGTG 2343
Db 1507 AAAGGAATTGGAGAAATGAACAAGTAGATAAATTAGTCAAGTGTGGATCCGGAAGGTG 1566
QY 2344 CTGTTCTCTGGACGGCATCGAT 2364
Db 1567 CTGTTCTCTGGACGGATCGAT 1587

RESULT 5

US-09-735-487-7
; Sequence 7, Application US/09735487
; Patent No. US20020042679A1
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/735,487
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/117,217
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(492)
; OTHER INFORMATION: gag Polyprotein
US-09-735-487-7

Query Match 44.9%; Score 1106.8; DB 10; Length 2601;
Best Local Similarity 68.2%; Pred. No. 8.7e-176;
Matches 1618; Conservative 0; Mismatches 732; Indels 22; Gaps 5;

QY 14 TGGCCGAGGCGCATGAGCAGCGCACCA--GCGCCAACTCCTGATGCGCGAGCAACT 70
Db 77 TGGCTGAAGCAATGAGCCAAAGTACAAATTCAGTCCATTAATGATGAGAGGCAATT 136
QY 71 TCAGAGGCCCCAAGCGGCATCATCAAGTCTTCAACTCGGCAAGGAGGGCCACATCGCC 130
Db 137 TTAGGAACCAAGAAAGATTGTTAAGTGTTCATTTCTGCAAAAGAGGGCACACGCCA 196
QY 131 GCAACTGCGCGCGCCCCCGCAAGAGGGCTGTGGAAGTGGGCAAGAGGGCCACACAGA 190
Db 197 GAAATTCAGGGGCCCCCTAGGAAAAGAGGGCTGTGGAATTTGGAAGAGAGGACACAAA 256
QY 191 TGAAGGACTGCACGCGCAGCGCCAACTTCTTCGCGAGGAGCTGGCTTCCCGCCAGG 250
Db 257 TGAAGATTTGCTACTGAGAGACAGGCTAA-TTTTATGGGAAGATCTGGCTTCTTACAAG 315
QY 251 GCAAGGCCCGCGAGTTCCCGAGGAGCAACCGCGCAACAGCCCGCCAGCGCGAGC 310
Db 316 GGAAGGCCAGGGAATTTCTTCAGAGCAGACACAGCCACAGCCCGCCAGAGAGAGC 375
QY 311 TGCAGGTGCGCGG-----CGAACCCCGCAGGAGCGCGCGCGAGCGCGAGGCA 364
Db 376 TTCAGGTCTGGGTAGAGACAACAACCTCCCGCTCAGAAAGCAGGAGCGGATAGACAAGAA 435
QY 365 -----CCCTGAAGTCTCCCGCAGATCACCTGTGGCAGCGCCCGCTGCTGAGCATCAAG 418
Db 436 CTGTATCTTAACTTCCCTCAGTCTACTTTTGGCAACGACCCCTCTCACATAAGA 495


```

; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/117,217
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 9

```

Query Match	44.9%;	Score 1106.8;	DB 10;	Length 2601;
Best Local Similarity	68.2%;	Pred. No. 8.7e-176;		
Matches 1618;	Conservative	0;	Mismatches 732;	Indels 22;
Gaps				
QY	14	TGCGCGAGGCCATGAGCGAGGCCACCA--GCGCCAACCTCTGATGCAGCGCAGCAACT	70	
DB	77	TGCTTGAAGCAATGAGCCAGTTAACAATTCAGCTTCCATAATGATGCAGAGGCGAATT	136	
QY	71	TCAAGSGCCCCAAGCGCATCATCAAGTGTCTTCAACTTCGSCAAAGGAGGCCACATCGCCC	130	
DB	137	TTAGGAACCAAGAAGATTGTTAAGTGTGTTTCAATTGTGCAAGAAGGSCACACAGCCA	196	
QY	131	GCAACTCGCCGCCCCCGCAAGAGGGCTGTGGAAAGTGCAGAGGAGGCCACACAGA	190	
DB	197	GNAATTTCAGGGCCCCCTAGAAAAGAGGGCTGTGGAAATGTGAAAAGAGACACCAAA	256	
QY	191	TGAAGGACTCCACGAGCGCCAGGCCAACTTCTTCGCGAGGAGACCTGGCCTTCCCCCAGG	250	
DB	257	TGAAGATTGTCTGAGAGACAGGCTTAA-TTTTTTATGGGAAGATCTGGCCCTTCTTACAAG	315	
QY	251	GCAAGGCCCGCGAGTTCCCGAGGAGCGAGAACCGCGCCACACGCCCCACCGACGCGGAGC	310	
DB	316	GGAAGGCCAGGGAAATTTCCTTCAGAGCAGACACAGAGCCAAACGCCCCACCAGAAGAGC	375	
QY	311	TGCAGGTGCGCGG-----CGACAACCCCGCAGCAGGCGCGCGCGCAGCGCCAGGGCA	364	
DB	376	TTCAGGFTCTGGGGTGTAGAGACAACAACTTCCCCCTCAGAAGCAGGAGCGATAGACAAGAA	435	
QY	365	-----CCCTTGAACTTCCCCCAGATCACCTGTGGCAGCGCCCCCTGGTGAGCATCAAGG	418	
DB	436	CTGTATCTTTAACTTCCCTCAGGTCACTCTTTGGCAACGACCCCTCGTCACAATTAAGA	495	
QY	419	TGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCCAGCACACCGTGTCTGGAGG	478	
DB	496	TAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGATTATGAAG	555	
QY	479	AGATGAGCCTGCCCGCAGTGAAGCCCCAAGATGATCGCGCGCATCGCGGGCTTCATCA	538	
DB	556	AAATAGTTTTGCCAGGAAGATGAAACCAAAATATAGGGGGAATTTGAGGTTTTTATCA	615	
QY	539	AGGTGCCCACTACGACCAATCCTGATCGAGATCTGCGSCAAGAAGGCCATCGGCACCG	598	
DB	616	AAGTAAGACAGTATGATCAGATCTCATAGAAATCTGTGGACATAAGCTATAGGTTACAG	675	
QY	599	TGCTGTATCGGCCCCACCCCGTGAACATCATTCGCGCGCAACATGCTGACCCAGCTGGGCT	658	
DB	676	TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTTCAGATTTGGTT	735	
QY	659	GCACCTGAACTTCCCATCAGCCCCATCGAGACCGTGGCCGTGAAGCTTGAAGCCCGGCA	718	
DB	736	GCACCTTTAAATTTTCCCATTTAGCCCTATTGAGACTCTAGCCAGTAAATTTAAAGCCAGAA	795	
QY	719	TGAGGSCCCCAAGGTGAAGCACTGGCCCTGACCGAGAGAGATCAAGGCCCTGACCG	778	
DB	796	TGATGGCCCCAAAGTTAAACAATGGCCATTGACAGAAGAAAATAAAGCAATTAGTAG	855	
QY	779	CCATCTCGGAGAGATGGGAAGAGGGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT	838	

Db 556 AAATGAGTTTCCAGGAAGATGGAACCAAAATGATAGGGGAATTTGGAGTTTATCA 615
QY 539 AGTGCCCGCAGTACGACCAAGTCTGTATGAGATCTCGGCGAAGAGCCATCGGACCG 598
Db 616 AAGTAAGACAGTATGATCAGATCTATAGAAATCTGTGACATATAAGCTATAGGTACAG 675
QY 599 TGCATGATCGGCCCGCCAGTGAACATCATCGGCCGCAACATCTGACCCAGCTGGGCT 658
Db 676 TATTAGTAGACCTACACCTGTCAACATATTGGAAGAAATCTGTTGACTCAGATTGGTT 735
QY 659 GCACCCCTGAATCTCCCATCAGCCCCATCGAGACCGTGCOCGTGAAGCTGAAGCCCGCA 718
Db 736 GCACCTTTAAATTTCCCATATAGCCCTATTGAGACTGTACAGTAAATTTAAAGCCAGAA 795
QY 719 TGGAGCGCCCGCAGGTGAAGCAGTGGCCCTTGACCGGAGGAGAAATCAAGCCCTGACC 778
Db 796 TGGATGCGCCCAAGATTAAGCAATTTGGCCATTGACAGAGAAATAAAGCATTAGTAG 855
QY 779 CCATCTCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCAGAACCCCT 838
Db 856 AAATTTGTACAGATGGAAGAGGAGGAAATTTCAAAATTTGGGCTGAAATCCAT 915
QY 839 ACAACACCCCGTTCGCCATCAAGAGAGGAGACACCAAGTGGCGCAAGCTGGTG 898
Db 916 ACAATCTCCAGTATTGGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAAATTAGTAG 975
QY 899 ACTTCGGGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGAGTGGGCATCCGCC 958
Db 976 ATTTTCAGAGAACTTAATTAAGAGAACTCAAGACTTCTGGGAAATTTCAATTAGGAATACCAC 1035
QY 959 ACCCGCGCGCTCAAGAAGAAAGAGCGTGCAGCGTCTGACGTGGGGCGACCGCTACT 1018
Db 1036 ATCCCGCAGGCTTAAAGAGAAATTAAGTAACTGACAGTACTGATGGGTGATGCATATT 1095
QY 1019 TCAGCGTCCCTTGAGCAGGAGTTCGCAAGTACACCGCTTCCACCATCCCGCAGCATCA 1078
Db 1096 TTTTCAGTTCCTTAGTGAAGACTTCAGGAAGTATAGTGCATTTACCATACTAGTATAA 1155
QY 1079 ACAAGGACCCCGGATCCGCTACCAAGTACACAGTGTCTCCCGCAGGGCTGGAAGGCA 1138
Db 1156 ACAATGAGACACCGAGGATAGATATAGTACATGATGCTTCCACAGGATGGAAAGAT 1215
QY 1139 GCGCCAGCATCTCCAGAGCAGTACCAAGATCTCGAGCCCTTCGCGCCCGCAGAC 1198
Db 1216 CACCAGCAATATTCAGAGTACATGACAAATCTTAGAGCCCTTTAGAAAACAAATC 1275
QY 1199 CCGAGATCGTATACCA-----GGCCCCCTGTAGCTGGGCGAGCAGCTGGAGATCG 1252
Db 1276 CAGACATAGTTTATCTCAATACATGATGATGTTGTATGTAGGATCTGACTTAGAAATAG 1335
QY 1253 GCCAGCAGCGCGCAAGATCCAGAGCTGCGCAAGCAGCTGCTGGCTGGGCTTCACCA 1312
Db 1336 GGCAGCATAGAACAAATAGAGGAGCTGAGACACATCTCTTGGAGTGGGAGCTTACCA 1395
QY 1313 CCCCGCAAGAAGCACAGAGGAGGCCCTTCCTGTGGATGGGCTACAGCTGCACC 1372
Db 1396 CACCAGACAAAACATCAGAAAGACCTCCATCTTGTGATGGTTTGAACCTCCATC 1455
QY 1373 CCGACAAGTGACCGTGCAGCCCATCGAGCTGCGCGAAGAGAGAGCTGGACCGTGAACG 1432
Db 1456 CTGATAATGGCAGTACAGCCTATAGTGTCTCCGAGAAAAGACAGCTGTCAATG 1515
QY 1433 ACATCCAGAAGCTGTGGGAGCTGAACTGGCCAGCAGCATCTACCCCGCATCAGG 1492
Db 1516 ACATACAGAAGTATAGTGGGAAATTTGAATTTGGCAAGTCAATTTACCAGGGATTAAG 1575
QY 1493 TCGCCAGCTGTGAAGCTGTGGCGCGCGCAAGCCCTCACCGACATCGTGCOCCTGA 1552
Db 1576 TAAGCAATTTATGTAACCTCTTAGAGGAACCAAGACACTAACAGATTAATACCCTAA 1635
QY 1553 CCGAGGAGCGCAGCTGGAGCTGGCCGAGAACCCGCGAGATCTCGCGGAGCCCGTGCACG 1612

Db 1636 CAGAAGAGCAGAGCTAGAACTGGCAGAAAAACAGAGAGATTCTAAAAAGAACCACTACATG 1695
QY 1613 GCCTGTACTAGCACCCAGCAGGAGCACTGGTGGCCGAGATCCAGAAGCAGGCGCACAGACC 1672
Db 1696 GAGTGTATTATGACCCCATCAAAAGACTTAATAGCAGAAATACAGAAAGCGGGCAAGGCC 1755
QY 1673 AGTGACCTACAGATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCCGCAAGTACGCCA 1732
Db 1756 AATGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGAAAATATGCAA 1815
QY 1733 AGATGGCGACCGCCACACCAACGACGTGAAGCAGCTGACCGAGCCCTGCAGAAAGATCG 1792
Db 1816 GAATGGGGGTGCCCACACTAATGATGTAAACAAATTTAACAGAGCAGTGCAAAATAA 1875
QY 1793 CCATGGAGAGCATCGTATCTGGGCAAGACCCCAAGTTCGCGCTGCCCTCCAGAAAGG 1852
Db 1876 CCACAGAAAGCATAGTATATGTTGGGAAAGACTCTTAATATTTAAACTGCCATACAAAAG 1935
QY 1853 AGACTGGGAGACCTTGTGTGACCGCACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1912
Db 1936 AAACATGGAAACATGTTGGACAGAGTATTGGCAAGCCACCTGGATTCTTGTGTTGGAGT 1995
QY 1913 TCGTGAACACCCCGCTGTGTGAAGCTGTGTACCACTGGAGAGAGAGCCCATCATCG 1972
Db 1996 TTGTTAATACCCCTCCCTTAGTGAATTTATGTTACAGTTTAGAGAAAAGAACCCCATAGTAG 2055
QY 1973 GCGCGGAGACCTTCTACGTGGACGCGCCGCAAGCTGACGAGTTCGCGAGACCAAGATCGCAAGGCG 2032
Db 2056 GAGCAGAAACCTTCTATGTAGTGGGCGAGCTAACAGGAGACTAAATTTAGAAAAGCAG 2115
QY 2033 GCTAGTGAACGACCGCGGCGGCAAGAGATCGTGAAGCTGACCGCTGACCGAGACCAACACCA 2092
Db 2116 GATATGTTACTAATAGAGAGACAAAAGTTGTCAACCTTAACGACACACAAATCAGA 2175
QY 2093 AGACGAGCTGCAGCCATCCAGCTGGCCCTGACGAGCAGCGGAGAGGTGACATCG 2152
Db 2176 AGACTGATTTACAAGCAATTTATCTAGCTTTGCGAGGATTCGGGATTTAGAAAGTAAACATAG 2235
QY 2153 TGACCGACAGCAGTACGCTGCGCATCTCCAGCGCCAGCCCGCAGAGCAGAGCG 2212
Db 2236 TACAGATCAATATGCAATTTAGGAATCTTCAAGACACACCAAGTCAAGTGAATCAG 2295
QY 2213 AGCTGGTGAACCAAGATCATCGAGCAGCTGATCAAGAAGAGAGGTGTACCTGAGCTGGG 2272
Db 2296 AGTTAGTCAATCAAAATAATAGACAGTTAATAAAAAAGGAAAAGCTATCTTGGCATGGG 2355
QY 2273 TGCCCGCCCAAGGCGATCGGCGCAAGCAGCAGTGCAGAGCTGGTGGACAGGCA 2332
Db 2356 TACCAGCACAAAGGAATTTGAGGAAATGAACAAAGTAGATAAATTAGTCAAGTGTGGAA 2415
QY 2333 TCCGCAAGGTGTGTCTCTGCGAGCGCATCGAT 2364
Db 2416 TCAGAAAGTACTATTTTAGATGGAATAGAT 2447

RESULT 9
US-09-999-183-1
; Sequence 1, Application US/09999183
; Patent No. US20020147169A1
; GENERAL INFORMATION:
; APPLICANT: MITROPHANOUS, et al
; TITLE OF INVENTION: In Vivo Selection Method
; FILE REFERENCE: 674523-2009
; CURRENT APPLICATION NUMBER: US/09/999,183
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: PCT/GB00/02136
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 9912965.2
; PRIOR FILING DATE: 1999-06-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: SeqWin99
; SEQ ID NO 1
; LENGTH: 4307

Db	1320	GAAAUUCAGCGCCCUAGAGAAAAGGGGUGUUGAAUUGGAAAGAGGACACCAAA	1379
QY	191	TGAAGACTGCACCGAGCGCCAGGCGCAACTTCTTCGCGGAGGAGCACTGGCCTTCGCCCAGG	250
Db	1380	UGAAAGAUUUGUACUAGAGACACAGGCUAA-UUUUUUUGGGAAGAUUCUGGCUUUCUUAACAAG	1438
QY	251	GCRAAGCCCGCGAGT-----TCCCNAGCG	274
Db	1439	GGAAAGCCGAGGGAUUUUUCUACAGACAGACCAGAGCCACAGCCCAACCAUUUUUUCAG	1498
QY	275	AGCAGAAACGCGCCCAACACCCACAGCGCGGAGCTGCAGTGGCGGG-----CGACA	328
Db	1499	AGCAGACCAGACACACCCACCAGAGAGACUUCAGGUCUGGGUAGAGACACA	1558
QY	329	ACCCCGCAGCAGGCGCGCGCGAGCGCCAGGCA-----CCCTGAACCTTCCCCGAGA	382
Db	1559	ACUCCCCUACAAAGCAGGAGCGGAGACAGGAACUGUAUCCUUUAACUUUCCUUCAGA	1618
QY	383	TCACCTGTGGCAGCGCCCTGTGTAGCATCAAGTGGCGCGCCAGATCAAGAGGAGGCC	442
Db	1619	UCACUCUUUGCAACGACCCUUCUACAUAAGAUAAGGGGGCAUAAGAAGAGCU	1678
QY	443	TGCTGACACCGCGCGCAGACACCGTGTGGAGGAGATGAGCCTGCCCGCAAGTGA	502
Db	1679	UAUUAAGUACAGGACAGAUAGACAGUAUAGAAGAAUAGUUGUUGCAGGAAGAUGA	1738
QY	503	AGCCCAAGATGATCGCGGCATCGCGGCTTCATCAAGTGGCGCGCCAGTACGACCATCC	562
Db	1739	ACCAAAAUAGUAGGGGAUUUGAGGDUUUAUCAAAGUAAGACAGUUAUCAGAUAC	1798
QY	563	TGATCAGAGATCTGCGCAAGAGGCATCGGCACCGTCTGATCGGCCCAACCCCGTGA	632
Db	1799	UCAUAGAAUUCUGGACAUAAAGCUAUAAGGUACAGUAUAUAGUAGGACCUACACCUCA	1858
QY	623	ACATCATCGGCCCAACATGCTGACCCAGCTGGGCTGCACCTGAACCTTCCCAATCAGCC	682
Db	1859	ACAUAUUGAAGAAUUCUGUUGACUCAGAUUGUUGACAUUUAAAUUUUCCCAUAAGCC	1918
QY	683	CCATCAGACCGTGCOCGTGAAGCTGAAGCCGGCATCGACGGCCCAAGGTGAAGCAGT	742
Db	1919	CUAUAGACUGUACCAAGUAUAUAAGCCAGGAUUGUAGGCCCAAGUUAACAUA	1978
QY	743	GGCCCTTGCCGAGGAGAATCAAGGCCCTGACCGCCATCTCGAGAGAGATGGAGAAG	802
Db	1979	GGCCAUGACAGAAGAAAAAUAAGCAUAUAGUAAGAAUUUUGUACAGAAUAGGAAAGG	2038
QY	803	AGGCAAGATCACCAAGTCGCGCCGAGACCCCTACACACCCCGTGTTCGCCATCA	862
Db	2039	AGGGAAAAUUUUGCGGCUUGAGAUCCAUAUACUUAUCCAGUAUUUUGCCAUAA	2098
QY	863	AGAAGAAGGACACACCAAGTGGCGCAAGCTTGGTGGACTTCCGCAGCTGAACAAGCGCA	922
Db	2099	AGAAAAAGACAGUACUAAUUGGAGAAAAUUAUAGUAUUUCAGAGAACUUAUAAGAGAA	2158
QY	923	CCAGAGCTTCTGGGAGTGCAGCTGGGCATCCCCACCCCGCGCGCTGNAGAGAGA	982
Db	2159	CUCAGACUUCUGGGAUGUUAUAGUAUAACCAUCCCGCAGGGGUAAAAAGAAAA	2218
QY	983	AGAGCGTGCAGCTGTGGACGTGGCGACGCGCTACTTCAGCGTGCCTCGACGAGGACT	1042
Db	2219	AUACAGUAACAGUACUUGUGUGGUAUGCAUUAUUUCAGUUCUUCUUAAGUAGAGACU	2278
QY	1043	TCCGCAAGTACACCGCTTACACATCCCGAGCATCAACAAGACACCCCGCGCATTCGCT	1102
Db	2279	UCAGGAAGUAUACUGCAUUUACCAUCCUAGUUAUAACAUAGACACACAGGAUUAUAGU	2338
QY	1103	ACCAGTACAACTGTGTCGCCCGCAGGCTGAAGGGCAGCCCGACAGATCTTCCAGAGCAGA	1162
Db	2339	UACAGUACAUGUGUUCACAGGAGUAGGAAGGAUACACAGCAUAUUAUCCAAAGUAGCA	2398
QY	1163	TGACCAAGATCTGTGGAGCCCTTCCGCGCCGCAACCCCGAGATCGTATCTACCA-----	1217
Db	2399	UGACAAAAUUUAGAGCCUUUUAAAAACAACAAUCCAGACAUAGUUAUUAUUAUACA	2458

Qy 623 ACATCATGGCGCCCAACATGCTACCCAGCTGGGCTGCACCTGAACCTTCCCATCAGCC 682
Db 1852 ACATAAATTTGAAGAAATCTGTTGACTCAGATTGGTTGCACCTTAATTTTCCCATAGCC 1911
Qy 683 CCATCGACACGTCGCTGAGCTGAAGCTGAAGCCGGATCGAGCGGCCCAAGGTGAAGCAGT 742
Db 1912 CTATTGACACTGTACCACTAAATTTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAAT 1971
Qy 743 GGCCCTCACCAGGAGAGAAATCAAGGCCCTGACCGCCCATCTGCGAGGAGTGGAGAGG 802
Db 1972 GGCCATTGACAGAGAGAAATTAAGCATTTAGTAGAAATTTGACAGAAATGGAAAAGG 2031
Qy 803 AGGCAAGATCACCAGATCGGCCCGGAGAACCCCTACAACACCCGCCCTGTTCGCCATCA 862
Db 2032 AAGGAAAAATTTCAAAATTTGGGCTGAGATCCATACAATATCCAGTATTTGCCATAA 2091
Qy 863 AGAAGAGGACAGACCAAGTTGGGCAAGCTGGTGAGCTTCCGGAGCTGAACAAAGCGCA 922
Db 2092 AGAAAAAGACAGTACTAAATGGAGAAATTTAGTAGATTTTCAGAGAACTTAATAGAGAA 2151
Qy 923 CCCAGGACTTCTGGAGGTGCGAGCTGGGCATCCGCCACCCCGCGGCTGGAAGAGA 982
Db 2152 CTCAGACTTCTGGAGTTCAATTTAGGAATACCACATCCCGCGAGGTTAAAAAGAAAA 2211
Qy 983 AGAGCGTGACCGTCTGGACCTGGGCGACGCTTACTTCAGCGTGGCCCTGGACGAGGACT 1042
Db 2212 AATCAGTAACAGTACTGGATGGGTGATGCATATTTTTCAGTTCCCTTAGTAGAGACT 2271
Qy 1043 TCCGCAAGTACACCGCTTACCATCCCGACGATCAACAAGACGACCCCGGCTATCCGCT 1102
Db 2272 TCAGGAATTAATGCAATTTACCATACCTAGTATAAACAATGAGACACACGAGGATTAGAT 2331
Qy 1103 ACCAGTCAACGCTGTGCCCGCGCTGGAAGCGCAGCCCGACGATCTCCAGAGCAGCA 1162
Db 2332 ATCAGTACAATGCTTCCACAGGATGGAAGGATCACCAAGATATTCCAAAGTAGCA 2391
Qy 1163 TGACAAAGATCCTGGAGCCCTTCGCGCCCGCAACCCGAGATCGTGATCTACCA - - - - 1217
Db 2392 TGACAAAAATCTTAGAGCTTTTAAAAAACAATCCAGACATAGTTATCTATCAATACA 2451
Qy 1218 -GGCCCCCTGTAGTGGGACGACCTGGAGATCGGCGACGACCGCCCAAGATCGAGG 1276
Db 2452 TGAACGATTTGTATGTAGGATCTGACTTAGAATAGGCGCATAGAACAAAAATAGAGG 2511
Qy 1277 AGCTGCGCAAGCCTGTGCTGGCTGGGCTTCACACCCCGCAAGAACACACAGAGG 1336
Db 2512 AGCTGAGACACATCTGTTAGGTGGGACTTACCACACGACAAAAAATCAGAAAG 2571
Qy 1337 AGCCCCCTTCTGTGATGGGCTACGAGCTGACCCCGCAAGTGGACCTGCGAGCCCA 1396
Db 2572 AACCTCCATTCCTTGGATGGGTTATGAATCCATCTGATATAATGGACATCAGACCTA 2631
Qy 1397 TCGAGCTGCCGAGAGAGAGCTGGACCTGACGACATCCAGATCGGAGCTGGTGGGCAAGC 1456
Db 2632 TAGTGCTGCCAGAAAAAGACAGCTGGACTGCAATGATACATAGAAGTTAGTGGGGAAT 2691
Qy 1457 TGAAGTGGGCGAGCAGATCTACCCCGCATCAAGGTGCGGCGCTGTGCAAGCTGCTGC 1516
Db 2692 TGAATCCGCAAGTCAATTTACCCAGGGATTAAGTAAGCAATTAATGTAATCCCTTA 2751
Qy 1517 GCGGCGCAAGGCCCTGACCGACATCTGCCCCCTGACCCGAGGCGGAGCTGGAGCTGG 1576
Db 2752 GAGGAACCAAGCACTAACAGAAGTAATACCATTACAGAGAAGACGAGCTAGAAGTGG 2811
Qy 1577 CCGAAGACCGGAGATCTTGGCGAGCCGCTGACGCGGTGTACTAGACCCCGCAGAGG 1636
Db 2812 CAGAAAAACAGAGATTTCTAAAAAGAACCGAGTACATGAGGTGTTATGACCCATCAAAA 2871
Qy 1637 ACCTGGTGGCGGAGATCCAGAGCAGGCGCAGCAGCTGAGCTTACAGATCTACAGG 1696
Db 2872 ACTTAATAGAGAAATACAGAGCAGGCGCAAGGCCAATGACATATCAAAATTTATCAAG 2931

Qy 1697 AGCCCTTCAAGAAACCTGAAGACCGCGCAAGTACGCCAAGATGCGACCGCCACACCAAGC 1756
Db 2932 AGCCATTTAAATCTGAAACACAGAAATATGCAAGATGAGGGTGGCCACACATAATG 2991
Qy 1757 AGCTGAAGCAGCTGACCGAGGCGCTGCAGAAAGATGCCATGGAGAGCATCGTGATCTGGG 1816
Db 2992 ATGTAAAAACAATTAACAGAGGCGAGTCAAAAAATAAACCAAGAAAGCATAGTAATATGG 3051
Qy 1817 GCAAGACCCCAAGTTCGCCCTGCCATCCAGAACGAGACCTGGGAGACCTGGTGACCG 1876
Db 3052 GAAGAGCTCTAAATTTAAACTTACCCATACAAAGAAACATGGGAAACATGGTGACAG 3111
Qy 1877 ACTACTGCGAGCCCACTGGATCCCGGAGTGGAGTTCTGTGAACACCCCGCCCTGGTCA 1936
Db 3112 AGTATTGCAAGCCACTGGATTCTGTAGTGGAGTTGTTAATACCCCTCTCTTAGTGA 3171
Qy 1937 AGCTGTGTACAGCTGGAGAGGAGCCATCATCGCGCCCGAGACCTTCTACGTGGAGC 1996
Db 3172 AATTATGTTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAACCTTCTATGTAGATG 3231
Qy 1997 GCGCCCAACCCGAGACCAAGATCGCGAGGCGGCTACGTGACCGACCGGGCGCGC 2056
Db 3232 GGGCAGCTAACAGGGAGACTAAATTTAGGAAAAAGCAGGATATGTTACTAACAAAGGAAGAC 3291
Qy 2057 AGAAGATCGTGAGCCTGACCGAGACCAACACAGAAAGCAGCTGCAGGCGCATCCAGC 2116
Db 3292 AAAAGTTGTCCCTTAATACACAAACAAATCAGAAAACTCAGTTACAAAGCAATTTATC 3351
Qy 2117 TGGCCCTCGAGACAGCGGCGAGGTGAACATCGTGACCGACAGCGCAGTACGCCCTGG 2176
Db 3352 TAGCTTTCAGGATTCAGGATTTAGAAAGTAAACATAGTAACAGACTCACAAATATGCATTAG 3411
Qy 2177 GCATCATCCAGGCCCGCCGACACAGAGGAGGAGGAGCTGTGTGAACAGATCATCGAGC 2236
Db 3412 GAATCATTTCAAGCACACACCAAGATAAAAGTGAATCAGAGTTAGTCAATCAAAATAATAGAGC 3471
Qy 2237 AGCTGATCAAGAAAGAGAGGTGTACCTGAGCTGGTGGTGGCCCGCCACAGGCGCATCGCG 2296
Db 3472 AGTTAATAAANAAGGAAAGGTCTATCTGCGATGGGTACCAGCACAAAGAAATTTGGAG 3531
Qy 2297 GCACGAGCAGATCGACAAGCTGTGTGACAAAGGCGATCGCAAGGTGTCTTCTCTGGAGC 2356
Db 3532 GAAATGAACAAGTAGATAAATTAGTCAGTGTCTGGAATCAGGAAATACTATTTTTAGATG 3591
Qy 2357 GCATCGAT 2364
Db 3592 GAATAGAT 3599

RESULT 15

US-09-798-675-5
; Sequence 5, Application US/09798675
; Patent No. US20020106798A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: HIV VACCINES
; FILE REFERENCE: E056 2020
; CURRENT APPLICATION NUMBER: US/09/798,675
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/186,364
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/251,083
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 9918
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Construct of vaccine vector pGAL and vaccine insert expressing
; OTHER INFORMATION: ade B HIV-1 Gag-Po
; NAME/KEY: misc_feature

LOCATION: (106)..(1641)
; OTHER INFORMATION: encodes proteins of viral inner core
; NAME/KEY: misc.feature
; LOCATION: (1401)..(3617)
; OTHER INFORMATION: encodes viral proteins but not integrase
; NAME/KEY: misc.feature
; LOCATION: (3708)..(5715)
; OTHER INFORMATION: regulates high-level production of HIV genes
; NAME/KEY: misc.feature
; LOCATION: (3847)..(5944)
; OTHER INFORMATION: encodes protein regulating the transfer of RNA to cytoplasm
; NAME/KEY: misc.feature
; LOCATION: (3939)..(4181)
; OTHER INFORMATION: gene participates in viral assembly and budding
; NAME/KEY: misc.feature
; LOCATION: (4099)..(4941)
; OTHER INFORMATION: encodes truncated form of viral coat protein
; NAME/KEY: misc.feature
; LOCATION: (6098)..(9918)
; OTHER INFORMATION: vaccine vector pGAL
US-09-798-675-5

Query Match 42.5%; Score 1048; DB 10; Length 9918;
Best Local Similarity 66.9%; Pred. No. 5.3e-166;
Matches 1610; Conservative 0; Mismatches 740; Indels 58; Gaps 6;
QY 14 TGGCCGAGCCATGAGCCAGG---CCACCAGCCCAACATCTCTGATCGAGCGCAGCAACT 70
DB 1193 TGGCTGAAGCAATGAGCCAAAGTAAACAAATACAGCTACCAATAATGATGCGAGAGGCAATT 1252
QY 71 TCAAGGGCCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGCGCACATGCCCC 130
DB 1253 TTAGGAACCAAGAAAGATGGTTAAGAGCTTCAATAGCGGCAAGAGGCGCACACAGCCA 1312
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGGTGTCTGGAAGTGTGCGGCAAGGAGGCGCACACAGA 190
DB 1313 GAAATTGTCAGGGCCCCCTAGGAAAGAGGGCAGCTGGAAAGAGGAAAGAGGACACCAAA 1372
QY 191 TGAAGGACTCACGAGCGCCAGCCCAACTTCTTCCGCGAGGACCTGGCCCTCCCGCAGG 250
DB 1373 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTATTAGGGAAGATCTGGCCCTTCTTACAG 1431
QY 251 GCAAGGCCCGCGAGT-----TCCCCAGCG 274
DB 1432 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGAGCCACAGCCCAACCATTTCTTCAG 1491
QY 275 AGCAGAACCCGCGCACAGCCCGCCAGCGCGAGCTGCGAGGTGCGGGG-----CGACA 328
DB 1492 AGCAGACCAAGGCCAACAGCCCCCACCAGAGAGAGCTTCAGGTCTGGGGTAGAGACAACA 1551
QY 329 ACCCGCGCAGCGAGCGCGCGCGAGCGCCAGGCGCA-----CCCTGAACCTTCCCCCAGA 382
DB 1552 ACTCCCGCTCAGAGCAGGAGGCGGATAGACAGAGAACTGTATCTTTAACTTCCCTCAGA 1611
QY 383 TCACCTGTGGCAGCGCCCTGTGTAGCATCAAGTGGCGGCGCCAGATCAAGGAGGCC 442
DB 1612 TCACTCTTTGGCAACGACCCCTCGTCACATAAAGATAGGGGGCACTAAAGAGAGCTC 1671
QY 443 TGGTGGACCGCGCGCGAGCACCGTGTCTGGAGGAGATGAGCCTGCGCGGCAAGTGA 502
DB 1672 TATTATACAGGAGCAGATGATACAGTATTAGAAGAAATGAGTTTGGCCAGGAAGATGA 1731
QY 503 AGCCCAAGATGATCGCGGATCGCGGCTTCATCAAGTGGCCAGTACGACAGATCC 562
DB 1732 AACCAAAATGATAGGGGAAATTTGAGGTTTATCAAGTAAAGACAGATGATGATCAGATAC 1791
QY 563 TGATCGAGATCTGGGCAAGAGGCCATCGGCACCGCTGTGTGCGGCCCGCCACCGCCGTGA 622
DB 1792 TCATAGAAATCTGTGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCA 1851
QY 623 ACATCATCGCGCGAGCATCTGACCCAGCTGGGCTGCGACCCCTGGAACCTCCCATCAGCC 682
DB 1852 ACATAATTGGAAGAAATCTGTTGACTCAGATGTTGTTGCACTTTAAATTTTCCCATTAGCC 1911

QY 683 CCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGGCATGGACGGCCCAAGGTGAAGCAGT 742
DB 1912 CTATTGAGACTGTACCAGTAAATTAAGCCAGGAATGATGCCCAAAAGTTAAACAAT 1971
QY 743 GGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTCGGAGGAGATGAGAGAG 802
DB 1972 GGCCATTGACAGAAGAAATAAAGCATTAGTAGAAATTTTACAGAAATGAAAGG 2031
QY 803 AGGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTACACACCCCGCTTCGCCCATCA 862
DB 2032 AAGGGAAAATTTCAAAATTTGGCCCTGAGAAATCCATACAATCTCCAGTATTTGCCATAA 2091
QY 863 AGAAGAGGACGACACCAAGTGGCGCAAGCTGGTGGACTTCCCGGAGCTCAACAAGCGCA 922
DB 2092 AGAAAAAGACAGTACTAATATGGAGAAAATTTAGTAGATTTTACAGAACTTTAATAAGAGAA 2151
QY 923 CCAGGACTTCTGGGAGGTGCACTGGGCATCCCGCCACCCCGCGGCTGGAAGAAGAGA 982
DB 2152 CTCAGGACTTCTGGGAAGTTCAATTTAGGAATACCAATCCCGAGGGTTAAAAAAGAAA 2211
QY 983 AGAGCGTGACCGTGCTGGAGCTGGGCGACCGCTACTTCAGCGTGGCCCTCGAGGAGACT 1042
DB 2212 AATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTATAGTGAAGACT 2271
QY 1043 TCCGCAAGTACACCGCTTCCACCATCCCGAGCATCAACAACAGAGACCCCGGATCCGCT 1102
DB 2272 TCAGGAAGTATATGCACTTTACCATACCTAGTATAAACATAGACACACAGGATTAGAT 2331
QY 1103 ACCAGTACAAAGCTGCTGCCCGAGGGTGGAAAGGCGACGCCCGAGATCTTCCAGAGCAGCA 1162
DB 2332 ATCAGTACAATGTCTTCCACAGGGATGGAAGGATCACCAGCAATATTTCAAAGTAGCA 2391
QY 1163 TGACCAAGATCTTGGAGCCCTTCCGCGCCGCAACCCCGAGATCGTGTATCAACA 1217
DB 2392 TGACAAAATCTTAGAGCCCTTTAAAAAACAAAATCCAGACATAGTTATCTATCAATACA 2451
QY 1218 -GGCCCCCTTACGTGGCGAGGACCTGGAGATCGGCCAGCACCGCGCCCAAGATTCGAGG 1276
DB 2452 TGAACGATTTGTATGATAGATCTGACTTTAGAAATAGGCGAGCATAGAACAAAATAGAGG 2511
QY 1277 AGTGGCAGAGCAGCTGCTGCGGTGGGGCTTACACCCCGCGAGCAAGAGACAGAGG 1336
DB 2512 AGCTGAGACAACTCTGTTGAGGTGGGGACTTACCACACAGACAAAAAATCAGAGAG 2571
QY 1337 AGCCCCCTTCTGTGATGGGTACGAGTGCACCCCGACAGTGCAGCTGCAGCCCA 1396
DB 2572 AACCTCCATTCTTGGATGGGTATGAACCTCCATCTGATAAATGGACAGTACAGCCTA 2631
QY 1397 TCGAGCTGCCCGAGAGAGAGTGGACCGTGAACGACATCCAGAAAGCTGTGGGCAAGC 1456
DB 2632 TAGTGTGCCAGAAAAGACAGCTGGACTGTCAATGACATACAGAACTTAGTGGGAAAT 2691
QY 1457 TGAATCTGGCGAGCAGATCTACCCCGCATCAAGGTGGCGCAGCTGTGCAAGCTGTGC 1516
DB 2692 TGAATACCGCAAGTCAAGTATTAACCGAGGATTAAGTAAAGCAATTTATGTAATCTCTTA 2751
QY 1517 GCGGCGCAGAGCCCTGACCGACATCTGCCCCCTGACCGAGGAGCGGAGCTGAGCTGG 1576
DB 2752 GAGGAACCAAGCACTAACAGAACTAATACCACTAAACAGAAAGAGAGAGTGAACCTGG 2811
QY 1577 CCAGAACCCCGAGATCTCGCGGAGCGCGTGCACGCGGTGTACTAGACGCCCAAGCAGG 1636
DB 2812 CAGAAAACAGAGATTTCTAAAGAACCACTACATGAGTGTATTATGACCCATCAAAAG 2871
QY 1637 ACCTGTGGCGAGATCCAGAGCAGGCGCCACGACCTAGTGGACCTACAGATCTACAGG 1696
DB 2872 ACTTAATAGCAGAAATACAGAAAGCAGGGGCAAGGCCAATGACATATCAATATTTATCAAG 2931
QY 1697 AGCCCTTCAAGAACCTCAAGACCGGCAAGTAGCCCAAGATGCGCACCCCGCCACACCAACG 1756
DB 2932 AGCCATTTAAAAATCTGAAAACAGGAAAAATATGCAAGAAATGAGGGGTGCCCACTAATG 2991

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 10:53:38 ; Search time 6233.78 Seconds
(without alignments)
11470.669 Million cell updates/sec

Title: US-09-610-313-32

Perfect score: 2457

Sequence: 1 gtcagcgcaccatggccga.....gggctagcaccgggaattc 2457

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1:	gb_ba.*	GenEmbl.*
2:	gb_htg.*	
3:	gb_in.*	
4:	gb_em.*	
5:	gb_ov.*	
6:	gb_pat.*	
7:	gb_ph.*	
8:	gb_pl.*	
9:	gb_pr.*	
10:	gb_ro.*	
11:	gb_sts.*	
12:	gb_sy.*	
13:	gb_un.*	
14:	gb_vi.*	
15:	em_ba.*	
16:	em_fun.*	
17:	em_hum.*	
18:	em_in.*	
19:	em_mu.*	
20:	em_om.*	
21:	em_or.*	
22:	em_ov.*	
23:	em_pat.*	
24:	em_ph.*	
25:	em_pl.*	
26:	em_ro.*	
27:	em_sts.*	
28:	em_un.*	
29:	em_vi.*	
30:	em_htg_hum.*	
31:	em_htg_inv.*	
32:	em_htg_other.*	
33:	em_htg_mus.*	
34:	em_htg_pln.*	
35:	em_htg_rdd.*	
36:	em_htg_mam.*	
37:	em_htg_vrt.*	
38:	em_sy.*	
39:	em_htgo_hum.*	
40:	em_htgo_mus.*	
41:	em_htgo_other.*	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	2457	100.0	2457	6	AX455916	Sequence
2	2436.2	99.2	2463	6	AX455915	Sequence
3	2415.4	98.3	2469	6	AX455914	Sequence
4	2000.8	81.4	9166	6	AX427930	Sequence
5	1992.4	81.1	9788	6	AX427936	Sequence
6	1978.4	80.5	9169	6	AX427931	Sequence
7	1975	80.4	9194	6	AX427926	Sequence
8	1973.4	80.3	9194	6	AX427925	Sequence
9	1973.4	80.3	12411	6	AX427927	Sequence
10	1970	80.2	9785	6	AX427938	Sequence
11	1966.6	80.0	9167	6	AX427933	Sequence
12	1966.6	80.0	9170	6	AX427928	Sequence
13	1966.6	80.0	9782	6	AX427935	Sequence
14	1966.6	80.0	9783	6	AX427934	Sequence
15	1966.6	80.0	9792	6	AX427932	Sequence
16	1965	80.0	9189	6	AX427921	Sequence
17	1962	79.9	4353	12	AF287353	Synthetic
18	1945.4	79.2	3015	6	AX455946	Sequence
19	1928.8	78.5	3009	6	AX455987	Sequence
20	1905.2	77.5	4352	12	AF287352	Synthetic
21	1878.6	76.5	1965	6	AX455952	Sequence
22	1859.8	75.7	2577	6	AX457088	Sequence
23	1852.2	75.4	4343	6	AX306428	Sequence
24	1848.2	75.2	4341	6	AX306429	Sequence
25	1838.6	74.8	1977	6	AX455954	Sequence
26	1825.6	74.3	7897	6	AX427923	Sequence
27	1772	72.1	3024	12	AF287355	Synthetic
28	1757.2	71.5	9407	6	AX427937	Sequence
29	1626.6	66.2	4307	6	AX019132	Sequence
30	1607.4	65.4	4307	6	AX035453	Sequence
31	1607.4	65.4	4307	6	AX056836	Sequence
32	1599.4	65.1	4307	6	AX283597	Sequence
33	1599.4	65.1	4307	6	AX283607	Sequence
34	1599.4	65.1	4327	6	AX035465	Sequence
35	1599.4	65.1	4353	6	AX035464	Sequence
36	1599.4	65.1	4642	6	AX035463	Sequence
37	1599.4	65.1	9772	6	AX283584	Sequence
38	1596	65.0	1668	6	AX455962	Sequence
39	1554.4	63.3	1680	6	AX455960	Sequence
40	1539.2	62.6	8366	6	AX180752	Sequence
41	1503.8	61.2	4338	6	AX180747	Sequence
42	1495	60.8	3981	6	AX306430	Sequence
43	1274	51.9	4288	6	AX149648	Sequence
44	1230.6	50.1	9010	14	AF110974	HIV-1 iso
45	1230.6	50.1	9010	14	AF110975	HIV-1 iso

ALIGNMENTS

RESULT 1					
AX455916	AX455916	2457 bp	DNA	linear	PAT 06-JUL-2002
LOCUS	Sequence 32 from Patent WO0204493.				
DEFINITION	AX455916				
ACCESSION	AX455916.1	GI:21714901			
VERSION					
KEYWORDS	synthetic construct.				
SOURCE	synthetic construct				
ORGANISM	artificial sequences.				
REFERENCE	1				
AUTHORS	zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.				
TITLE	Polynucleotides encoding antigenic hiv type c polypeptides,				
JOURNAL	polypeptides and uses thereof				
	Patent: WO 0204493-A 32 17-JAN-2002;				

CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

Location/Qualifiers

1. -2457

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="PR975YMM"

BASE COUNT 566 a 837 c 754 g 300 t

ORIGIN

Query Match 100.0%; Score 2457; DB 6; Length 2457;

Best Local Similarity 100.0%; Pred. No. 3.1e-248; Indels 0; Gaps 0;

Matches 2457; Conservative 0; Mismatches 0; Gaps 0;

QY	1	GTGAGCGCCACCATGTCGCGAGGCGCATGAGCCAGGCGCCACACAGCGCCCAACATCTCTGATGCGAG	60
DB	1	GTGAGCGCCACCATGTCGCGAGGCGCATGAGCCAGGCGCCACACAGCGCCCAACATCTCTGATGCGAG	60
QY	61	CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTTCACATGCGGCGCAAGGAGGC	120
DB	61	CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTTCACATGCGGCGCAAGGAGGC	120
QY	121	CACATCGCGCAACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGCAGGCAAGGAG	180
DB	121	CACATCGCGCAACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGCAGGCAAGGAG	180
QY	181	GGCCACCATGATGAAGACTGCACCGAGCGCGCAGGCCAACTTCTTCGCGAGGACCTGGCC	240
DB	181	GGCCACCATGATGAAGACTGCACCGAGCGCGCAGGCCAACTTCTTCGCGAGGACCTGGCC	240
QY	241	TTCCCGCAGGCGAAGCGCGCGAGTTCGCCAGGCGAGCAGAACCGCGCCCAACAGCCCCACC	300
DB	241	TTCCCGCAGGCGAAGCGCGCGAGTTCGCCAGGCGAGCAGAACCGCGCCCAACAGCCCCACC	300
QY	301	AGCGCGAGCTGCGAGTGGCGGCGGACACACCCCGCAGGAGCGCGCGCGCGAGCGCCAG	360
DB	301	AGCGCGAGCTGCGAGTGGCGGCGGACACACCCCGCAGGAGCGCGCGCGCGAGCGCCAG	360
QY	361	GGCACCTGAACTTCCCGCAGATFACCCCTGTGGCAGCGCCCGCCCTGGTGAGCATCAAGGTTG	420
DB	361	GGCACCTGAACTTCCCGCAGATFACCCCTGTGGCAGCGCCCGCCCTGGTGAGCATCAAGGTTG	420
QY	421	GGCGGCGAGATCAAGGAGGCGCTTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCGAG	480
DB	421	GGCGGCGAGATCAAGGAGGCGCTTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCGAG	480
QY	481	ATGAGCTGCGCGCAAGTGGAAAGCCCAAGATCATCGCGGCGATCGCGGCGCTTCATCAAG	540
DB	481	ATGAGCTGCGCGCAAGTGGAAAGCCCAAGATCATCGCGGCGATCGCGGCGCTTCATCAAG	540
QY	541	GTGCGGCGAGTACGACAGATFCTGTGATCGAGATCTGGGCAAGAGGCCATCGGCAACCGTG	600
DB	541	GTGCGGCGAGTACGACAGATFCTGTGATCGAGATCTGGGCAAGAGGCCATCGGCAACCGTG	600
QY	601	CTGATCGGCGCCACCCCGTGAACATCATCGGCGGCAACATGCTGACCCAGCTGGGCTGC	660
DB	601	CTGATCGGCGCCACCCCGTGAACATCATCGGCGGCAACATGCTGACCCAGCTGGGCTGC	660
QY	661	ACCTGAACTTCCCGCATGAGCGCCATCGAGACCGTGGCGGCAAGCTGAAGCTGAAGCGCGCATG	720
DB	661	ACCTGAACTTCCCGCATGAGCGCCATCGAGACCGTGGCGGCAAGCTGAAGCTGAAGCGCGCATG	720
QY	721	GAGGCGCCCAAGGTGAAGCAGTGGCCCGCTGACCGAGGAGAGATCAAGGCGCTGACCGCC	780
DB	721	GAGGCGCCCAAGGTGAAGCAGTGGCCCGCTGACCGAGGAGAGATCAAGGCGCTGACCGCC	780
QY	781	ATCTGCGGAGAGTGAAGGAGGCGCAGATCACCAGATCGGCGCCCGCCCGAGACCCCTAC	840
DB	781	ATCTGCGGAGAGTGAAGGAGGCGCAGATCACCAGATCGGCGCCCGCCCGAGACCCCTAC	840
QY	841	AACACCCCGCTGTTCGCCATCAAGAGAGGAGCAGCACCAGTGGCGCAAGCTGGTGGAC	900
DB	841	AACACCCCGCTGTTCGCCATCAAGAGAGGAGCAGCACCAGTGGCGCAAGCTGGTGGAC	900

QY	901	TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCGATCCCCCAC	960
DB	901	TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCGATCCCCCAC	960
QY	961	CCGCGCGGCTGAAGAAGAAGAGAGCGTGAACCGTGTGACGTGGGCGAGCGCTACTTC	1020
DB	961	CCGCGCGGCTGAAGAAGAAGAGAGCGTGAACCGTGTGACGTGGGCGAGCGCTACTTC	1020
QY	1021	AGGCTGCCCTTGACGAGGACTTCCGCAAGTACACCGCTTACACCTCCACATCCCGACATCAAC	1080
DB	1021	AGGCTGCCCTTGACGAGGACTTCCGCAAGTACACCGCTTACACCTCCACATCCCGACATCAAC	1080
QY	1081	AACGAGACCCCGGCTACCGCTACCAAGTACAACGTCTGCTGCCAGGCGTGAAGGCGAGC	1140
DB	1081	AACGAGACCCCGGCTACCGCTACCAAGTACAACGTCTGCTGCCAGGCGTGAAGGCGAGC	1140
QY	1141	CCGAGATCTTCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACCCC	1200
DB	1141	CCGAGATCTTCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACCCC	1200
QY	1201	GAGATCTGATCTACCAAGCGCCCTGTACGTGGCGAGCGACCTGGAGATCGGCCAGCAC	1260
DB	1201	GAGATCTGATCTACCAAGCGCCCTGTACGTGGCGAGCGACCTGGAGATCGGCCAGCAC	1260
QY	1261	CGCGCAAGTGAAGAGTGGCGAAGCAGCTCTGCTGGTGGGCTTCAACACCCCGGAC	1320
DB	1261	CGCGCAAGTGAAGAGTGGCGAAGCAGCTCTGCTGGTGGGCTTCAACACCCCGGAC	1320
QY	1321	AAGAAGCACCAGAGGAGCGCCCTTCTGCCATGACGTGACCGCGACCCCGACAGTGGACC	1380
DB	1321	AAGAAGCACCAGAGGAGCGCCCTTCTGCCATGACGTGACCGCGACCCCGACAGTGGACC	1380
QY	1381	GTGAGCGCTTCAAGCTGCGCGAAGAGGAGTGGACCTGAACCTGAACACATCCAGAGCTG	1440
DB	1381	GTGAGCGCTTCAAGCTGCGCGAAGAGGAGTGGACCTGAACCTGAACACATCCAGAGCTG	1440
QY	1441	GTGGGCAAGCTGAACCTGGGCGCAGCATATCCCGGCGATCAAGGTGCGGCGCTGTGC	1500
DB	1441	GTGGGCAAGCTGAACCTGGGCGCAGCATATCCCGGCGATCAAGGTGCGGCGCTGTGC	1500
QY	1501	AAGCTCTGCGCGCGCCAGGCGCTGACCGAGATGCTGCCCTGACCGAGGAGGCGGAG	1560
DB	1501	AAGCTCTGCGCGCGCCAGGCGCTGACCGAGATGCTGCCCTGACCGAGGAGGCGGAG	1560
QY	1561	CTGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCGCTGACCGCGCTGTACTACGAC	1620
DB	1561	CTGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCGCTGACCGCGCTGTACTACGAC	1620
QY	1621	CCGAGCAAGACCTGGTGGCGGAGATCCAGAACGAGGCGCACACAGTGGACCTACCAG	1680
DB	1621	CCGAGCAAGACCTGGTGGCGGAGATCCAGAACGAGGCGCACACAGTGGACCTACCAG	1680
QY	1681	ATCTACGAGGAGCTTCAAGAACCTGAACCGCGCAAGTACGCCAAGATCGCGACCGCC	1740
DB	1681	ATCTACGAGGAGCTTCAAGAACCTGAACCGCGCAAGTACGCCAAGATCGCGACCGCC	1740
QY	1741	CACACCAACGCTGAAGCAGCTGACCGAGGCGCTGCGAAGATCGCCATGCGAGAGCATC	1800
DB	1741	CACACCAACGCTGAAGCAGCTGACCGAGGCGCTGCGAAGATCGCCATGCGAGAGCATC	1800
QY	1801	GTGATCTGGGCGAAGACCCCGGAGTTCGCGCTGCCATCCAGAGGAGACTGGGAGACC	1860
DB	1801	GTGATCTGGGCGAAGACCCCGGAGTTCGCGCTGCCATCCAGAGGAGACTGGGAGACC	1860
QY	1861	TGGTGAAGCCACTACTTGGCAGGCGACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCC	1920
DB	1861	TGGTGAAGCCACTACTTGGCAGGCGACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCC	1920
QY	1921	CCCTTGGTGAAGTGTGGTACCAAGCTGGAGAAGAGCCCATCATCGCGCGCGGAGACCTTC	1980
DB	1921	CCCTTGGTGAAGTGTGGTACCAAGCTGGAGAAGAGCCCATCATCGCGCGCGGAGACCTTC	1980
QY	1981	TACGTGGACGCGCGCGCCCAACCGGAGACCAAGATCGGCAAGCGCGCTAGTGACCGAC	2040

```
Db 1981 TACGTGGACGGCGCGGCAACCGGAGACCAAGATCGCAAGGCGGCTAGCTGACCGAC 2040
Qy 2041 CGGGCGCGGCAAGAGATCGTGAAGCTGACCGAGACCAACACCAAGAGACCGAGCTGCAG 2100
Db 2041 CGGGCGCGGCAAGAGATCGTGAAGCTGACCGAGACCAACACCAAGAGACCGAGCTGCAG 2100
Qy 2101 GCCATCAGCTGGCCCTGACGAGCAGCGGCGAGCGAGGTGAACATCGTGAACGACGCCAG 2160
Db 2101 GCCATCAGCTGGCCCTGACGAGCAGCGGCGAGCGAGGTGAACATCGTGAACGACGCCAG 2160
Qy 2161 TACGCGCTGGGCATCATCCAGGCGCCAGCCGACACAGAGCGAGAGCGAGCTGGTGAACGAG 2220
Db 2161 TACGCGCTGGGCATCATCCAGGCGCCAGCCGACACAGAGCGAGAGCGAGCTGGTGAACGAG 2220
Qy 2221 ATCATCAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGTGGCGCCGCCACAAG 2280
Db 2221 ATCATCAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGTGGCGCCGCCACAAG 2280
Qy 2281 GGCATCGGCGCAACGAGCAGATCGACAAGCTGGTGAAGGCGATCGCGCAAGGTGCTG 2340
Db 2281 GGCATCGGCGCAACGAGCAGATCGACAAGCTGGTGAAGGCGATCGCGCAAGGTGCTG 2340
Qy 2341 TTCTGTGACGCGCATCGATGGCGGCGATCGTGATCTACCAAGTACATGGAGCGCTGTACGTG 2400
Db 2341 TTCTGTGACGCGCATCGATGGCGGCGATCGTGATCTACCAAGTACATGGAGCGCTGTACGTG 2400
Qy 2401 GGCAGCGCGGCCCTAGATGATGATTAAGGCTTCCCGGGCTAGCACCGGTGAATTC 2457
Db 2401 GGCAGCGCGGCCCTAGATGATGATTAAGGCTTCCCGGGCTAGCACCGGTGAATTC 2457

RESULT 2
AX455915
LOCUS AX455915 2463 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 31 from Patent WO0204493.
ACCESSION AX455915
VERSION AX455915.1 GI:21714900
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial construct.
REFERENCE
1.
AUTHORS zur Megede, J., Barnett, S.W., Engelbrecht, S., and van Rensburg, E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof
JOURNAL Patent: WO 0204493-A 31 17-JAN-2002;
CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES
source
1. .2463
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="PR975YM"
BASE COUNT 567 a 835 c 759 g 302 t
ORIGIN
Query Match 99.2%; Score 2436.2; DB 6; Length 2463;
Best Local Similarity 99.6%; Pred. No. 4.6e-246;
Matches 2454; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

Qy 1 GTCGACGCCACCATGGCGGAGGCCATGAGCCAGGCCACCAAGGCTTCACTGGGCAAGAGGCG 120
Db 1 GTCGACGCCACCATGGCGGAGGCCATGAGCCAGGCCACCAAGGCTTCACTGGGCAAGAGGCG 120
Qy 61 CGCAGCAACTTCAGGCGCCCAAGCGCATCATCAAGTGTCTCACTGGGCAAGAGGCG 120
Db 61 CGCAGCAACTTCAGGCGCCCAAGCGCATCATCAAGTGTCTCACTGGGCAAGAGGCG 120
Qy 121 CACATCGCCGCAACTGCGCGGCCCGCCGCAAGAGGCTCTGTAAGTGGCGCAAGGAG 180
Db 121 CACATCGCCGCAACTGCGCGGCCCGCCGCAAGAGGCTCTGTAAGTGGCGCAAGGAG 180
Qy 181 GGCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCC 240
```

```
Db 181 GGGCACCAAGATGAAGGACTGCACCGAGCGCAGGCAACTTCTTCCGCGAGGACCTGGCC 240
Qy 241 TTCCCCCAAGGCAAGGCGCGGAGTTCCCAAGCGAGAGAACCGCGCAACAGCCCCACC 300
Db 241 TTCCCCCAAGGCAAGGCGCGGAGTTCCCAAGCGAGAGAACCGCGCAACAGCCCCACC 300
Qy 301 AGCGCGAGCTGCAGTTCGGCGGCGACAAACCCCGAGCGAGCGCGCGCGCGCGAGCCAG 360
Db 301 AGCGCGAGCTGCAGTTCGGCGGCGACAAACCCCGAGCGAGCGCGCGCGCGCGAGCCAG 360
Qy 361 GGCACCTGAACTTCCCCCAAGATCAACCTGTGGAGCGCCCGCTGTGAGCATCAAGGTG 420
Db 361 GGCACCTGAACTTCCCCCAAGATCAACCTGTGGAGCGCCCGCTGTGAGCATCAAGGTG 420
Qy 421 GGCGCCAGATCAAGGAGCGCTGTGGACACCGCGCGGAGCACCATGCTGTGGAGGAG 480
Db 421 GGCGCCAGATCAAGGAGCGCTGTGGACACCGCGCGGAGCACCATGCTGTGGAGGAG 480
Qy 481 ATGAGCCTGCCCGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGGCTTCATCAAG 540
Db 481 ATGAGCCTGCCCGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGGCTTCATCAAG 540
Qy 541 GTGCGCCAGTACGACCCAGATCCTGATCGAGATCTCGGCAAGAGGCCATCGGCAACCGTG 600
Db 541 GTGCGCCAGTACGACCCAGATCCTGATCGAGATCTCGGCAAGAGGCCATCGGCAACCGTG 600
Qy 601 CTGATCGCCCGCCCGCGCGTGAACATCATCGCGCGACACATGCTGACCGAGCTGGGCTGC 660
Db 601 CTGATCGCGCCCGCCCGCGTGAACATCATCGCGCGACACATGCTGACCGAGCTGGGCTGC 660
Qy 661 ACCCTGAACTTCCCGCATCAGCCCCATCGAGACCGCGCGCTGAAGCTGAAGCCCGGCGATG 720
Db 661 ACCCTGAACTTCCCGCATCAGCCCCATCGAGACCGCGCGCTGAAGCTGAAGCCCGGCGATG 720
Qy 721 GACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCC 780
Db 721 GACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCC 780
Qy 781 ATCTGCGAGGAGATGGAGAGGAGGCGCAAGATCACCAGATCGGCGCGAGAACCCCTAC 840
Db 781 ATCTGCGAGGAGATGGAGAGGAGGCGCAAGATCACCAGATCGGCGCGAGAACCCCTAC 840
Qy 841 AACACCCCGCTGTTCCGCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGTGGAC 900
Db 841 AACACCCCGCTGTTCCGCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGTGGAC 900
Qy 901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGATCCCCAC 960
Db 901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGATCCCCAC 960
Qy 961 CCGCGCGCTGAGAGAGAGAGAGCGCTGACCGTGTGGAGCGCGCTGGGAGCGCTTACTTC 1020
Db 961 CCGCGCGCTGAGAGAGAGAGAGAGCGCTGACCGTGTGGAGCGCGCTGGGAGCGCTTACTTC 1020
Qy 1021 AGCGTGGCCCTGGAGGAGCTTCCGCAAGTACACCGCTTCAACATCCCGAGCATCAAC 1080
Db 1021 AGCGTGGCCCTGGAGGAGCTTCCGCAAGTACACCGCTTCAACATCCCGAGCATCAAC 1080
Qy 1081 AACGAGACCCCGCGCATCCGCTACAGTACAACTGTCTGCCCGCGAGGCTGGAAGGCGAGC 1140
Db 1081 AACGAGACCCCGCGCATCCGCTACAGTACAACTGTCTGCCCGCGAGGCTGGAAGGCGAGC 1140
Qy 1141 CCCAGCATCTTCCAGAGCAGATGACCAAGATCTGTGGAGCGCTTCCGCGCGCGCAACCC 1200
Db 1141 CCCAGCATCTTCCAGAGCAGATGACCAAGATCTGTGGAGCGCTTCCGCGCGCGCAACCC 1200
Qy 1201 GAGATCGTGTACTACGAGCGCCCTGTACGTGGCAGCGACCTGGAGATGGCGCAGCAG 1260
Db 1201 GAGATCGTGTACTACGAGCGCCCTGTACGTGGCAGCGACCTGGAGATGGCGCAGCAG 1260
Qy 1261 GCGGCCAAGATCGAGGAGCTGCGCAAGCACCTGTGCGCTGGGCTTCACCAACCCCGCAC 1320
Db 1261 GCGGCCAAGATCGAGGAGCTGCGCAAGCACCTGTGCGCTGGGCTTCACCAACCCCGCAC 1320
```

```
Db 1261 CGGCCAAGATCGAGGAGCTGCCAAGACCTCTGCTGGCTGGGGCTTACCACCCCGGAC 1320
QY 1321 AAGAAGCACAGAGGAGGCCCCCTTCTCCCTCCAT-----CGAGCTGACCCCGACAAG 1374
Db 1321 AAGAAGCACAGAGGAGGCCCCCTTCTTCTGGATGGCTACGAGCTGCACCCCGACAAG 1380
QY 1375 TGGACCGTGCAGCCCATCGAGCTGCGCGGAGAGAGAGAGAGTGGACCGTGAACGACATCCAG 1434
Db 1381 TGGACCGTGCAGCCCATCGAGCTGCGCGGAGAGAGAGAGTGGACCGTGAACGACATCCAG 1440
QY 1435 AAGCTGTGGGCAAGTGAACCTGGCGGACGACATCTACCCCGCATCAAGGTGCCCGAG 1494
Db 1441 AAGCTGTGGGCAAGTGAACCTGGCGGACGACATCTACCCCGCATCAAGGTGCCCGAG 1500
QY 1495 CTGTGAAGTGTCTGGCGGCGGCAAGGCGCTGACCGACATCTGCGCCCTGACCCGAGGAG 1554
Db 1501 CTGTGAAGTGTCTGGCGGCGGCAAGGCGCTGACCGACATCTGCGCCCTGACCCGAGGAG 1560
QY 1555 GCCGAGCTGGAGTGGCGCGGAGAACCGCGAGATCCTGCGGAGGCCCTGACGCGGTGTAC 1614
Db 1561 GCCGAGCTGGAGTGGCGCGGAGAACCGCGAGATCCTGCGGAGGCCCTGACGCGGTGTAC 1620
QY 1615 TACGACCCAGAGGAGACCTGTGGCGGAGATCCAGAACGAGGCGCACGACCATGTGACC 1674
Db 1621 TACGACCCAGAGGAGACCTGTGGCGGAGATCCAGAACGAGGCGCACGACCATGTGACC 1680
QY 1675 TACCAGATCTTACCAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1734
Db 1681 TACCAGATCTTACCAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1740
QY 1735 ACCGCCCCACACAGAGTGAAGAGCTGACGAGGCGCGTGAGAGATCGCCATGGAG 1794
Db 1741 ACCGCCCCACACAGAGTGAAGAGCTGACGAGGCGCGTGAGAGATCGCCATGGAG 1800
QY 1795 AGCATGCTGATCTGGGCAAGACCCCAAGTTCCTCGCTGCCATCCAGAGGAGACCTGG 1854
Db 1801 AGCATGCTGATCTGGGCAAGACCCCAAGTTCCTCGCTGCCATCCAGAGGAGACCTGG 1860
QY 1855 GAGACCTGTGTGACGACTACTTGGCAGGCGCACCTGGATCCCGAGTGGGAGTTCGTGAAC 1914
Db 1861 GAGACCTGTGTGACGACTACTTGGCAGGCGCACCTGGATCCCGAGTGGGAGTTCGTGAAC 1920
QY 1915 ACCGCCCTCTGTGAAGCTGTGGTACCAGCTGGAGAGAGGCCCATCATCTGCGCGCGGAG 1974
Db 1921 ACCGCCCTCTGTGAAGCTGTGGTACCAGCTGGAGAGAGGCCCATCATCTGCGCGCGGAG 1980
QY 1975 ACCTTCTACGTGACGCGCGCCCAACCGGACACCAAGATCGCAAGGCGCGCTACGCTG 2034
Db 1981 ACCTTCTACGTGACGCGCGCCCAACCGGAGACCAAGATCGCAAGGCGCGCTACGCTG 2040
QY 2035 ACCGACCGGGCGCGGAGAGATCGTGAGCCTGACCGGACACCAACCAAGAGACCGAG 2094
Db 2041 ACCGACCGGGCGCGGAGAGATCGTGAGCCTGACCGGACACCAACCAAGAGACCGAG 2100
QY 2095 CTGAGGCCATCCAGCTTGGCCCTGCAAGGACAGCGGAGAGGTGAACATCTGTGACCCGAC 2154
Db 2101 CTGAGGCCATCCAGCTTGGCCCTGCAAGGACAGCGGAGAGGTGAACATCTGTGACCCGAC 2160
QY 2155 AGCCAGTACGCCCTTGGCATCATCCAGGCCACCGCCGACAGAGCGAGCGAGCTGGTG 2214
Db 2161 AGCCAGTACGCCCTTGGCATCATCCAGGCCACCGCCGACAGAGCGAGCGAGCTGGTG 2220
QY 2215 AACGAGATCATCGAGCAGCTGTATCAAGAGGAGAGGTGTACCTGAGCTGGGTGCCCGCC 2274
Db 2221 AACGAGATCATCGAGCAGCTGTATCAAGAGGAGAGGTGTACCTGAGCTGGGTGCCCGCC 2280
QY 2275 CACAAGGCATCGCGGCAACGAGCAGATCGACAAGCTGGTGAAGAGGCGCATCCGCAAG 2334
Db 2281 CACAAGGCATCGCGGCAACGAGCAGATCGACAAGCTGGTGAAGAGGCGCATCCGCAAG 2340
QY 2335 GTGCTGTCTGACGCGCATCGATGCGGCGCATCTGATCTACCAAGTACATGGACGACCTG 2394
Db 2341 GTGCTGTCTGACGCGCATCGATGCGGCGCATCTGATCTACCAAGTACATGGACGACCTG 2400

QY 2395 TACGTGGCGACGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACCCGTTGAA 2454
Db 2401 TACGTGGCGACGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACCCGTTGAA 2460
QY 2455 TTC 2457
Db 2461 TTC 2463

RESULT 3
AX455914 AX455914 2469 bp DNA linear PAT 06-JUL-2002
LOCUS Sequence 30 from Patent WO0204493.
DEFINITION
ACCESSION AX455914
VERSION AX455914.1 GI:21714899
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
ARTIFICIAL SEQUENCES.
REFERENCE
1
AUTHORS zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL Polypeptides and uses thereof
PATENT: WO 0204493-A 30 17-JAN-2002;
CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES
Location/Qualifiers
1..2469
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="PR975(+)"
BASE COUNT 571 a 833 c 761 g 304 t
ORIGIN
Query Match 98.3%; Score 2415.4; DB 6; Length 2469;
Best Local Similarity 99.3%; Pred. No. 6.9e-244;
Matches 2451; Conservative 0; Mismatches 6; Indels 12; Gaps 2;
QY 1 GTCGAGCCACCATGGCGGAGGCGCATGAGCGCCACCGAGCCACCGCCACATCTCTGATGCAG 60
Db 1 GTCGAGCCACCATGGCGGAGGCGCATGAGCGCCACCGAGCCACCGCCACATCTCTGATGCAG 60
QY 61 CGCAGCAACTTCAAGGGCCCAAGCGCGATCATCAAGTGTTCACATCGCGCAAGGAGGCG 120
Db 61 CGCAGCAACTTCAAGGGCCCAAGCGCGATCATCAAGTGTTCACATCGCGCAAGGAGGCG 120
QY 121 CACATCGCGCGCAACTGCGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGGCAAGAG 180
Db 121 CACATCGCGCGCAACTGCGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGGCAAGAG 180
QY 181 GGCACACAGATGAAGGACTGCACCGAGCGCGAGCCAACTTCTCCGCGAGGACCTGGCC 240
Db 181 GGCACACAGATGAAGGACTGCACCGAGCGCGAGCCAACTTCTCCGCGAGGACCTGGCC 240
QY 241 TTCCCCCAGGGCAAGGCCCGCGAGTTCCCGAGCGAGCAGAACCCGCGCAACAGCCCCACC 300
Db 241 TTCCCCCAGGGCAAGGCCCGCGAGTTCCCGAGCGAGCAGAACCCGCGCAACAGCCCCACC 300
QY 301 AGCCGGAGCTGAGGTGCGCGGCGACACCCCGCGAGCGCGCGCGCGCGCGCGCGCGAG 360
Db 301 AGCCGGAGCTGAGGTGCGCGGCGACACCCCGCGAGCGCGCGCGCGCGCGCGCGCGAG 360
QY 361 GGCACCTGAACCTTCCCGCAGATCACCTCTGTGCGAGCGCCCTTGGTGAAGTCAAGGTG 420
Db 361 GGCACCTGAACCTTCCCGCAGATCACCTCTGTGCGAGCGCCCTTGGTGAAGTCAAGGTG 420
QY 421 GCGCGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGAG 480
Db 421 GCGCGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGAG 480
QY 481 ATGAGCCTGCCCGGCAAGTGGAGCGCCAGATGATGCGCGCGCATCGCGCGCTTCATCAAG 540
Db 481 ATGAGCCTGCCCGGCAAGTGGAGCGCCAGATGATGCGCGCGCATCGCGCGCTTCATCAAG 540
```

Qy	541	GTGCGCCAGTACGACCAGATCCTGATCGAGATCTCTGCGCAAGAAAGGCCATCGGCACCGTG	600
Db	541	GTGCGCAGTACGACCAGATCCTGATCGAGATCTGCGCAAGAAAGGCCATCGGCACCGTG	600
Qy	601	CTGATCGGCCCCACCCCGTGAACATCATCGCGGCCACATGCTGACCCAGCTGGCGTGC	660
Db	601	CTGATCGGCCCCACCCCGTGAACATCATCGCGGCCAACATGCTGACCCAGCTGGCGTGC	660
Qy	661	ACCTGAACTTCCCATCAGCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCATG	720
Db	661	ACCTGAACTTCCCATCAGCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCATG	720
Qy	721	GAGCGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC	780
Db	721	GAGCGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC	780
Qy	781	ATCTGCGAGGAGATCGAAGAGGCGCAGATCACCAGATCGGCCCGGAGAACCCCTAC	840
Db	781	ATCTGCGAGGAGATCGAAGAGGCGCAGATCACCAGATCGGCCCGGAGAACCCCTAC	840
Qy	841	AACACCCCGCTGTTCCGCATCAAGAAGAGGACAGACCAAGTGGCGCAAGCTGGTGGAC	900
Db	841	AACACCCCGCTGTTCCGCATCAAGAAGAGGACAGACCAAGTGGCGCAAGCTGGTGGAC	900
Qy	901	TTCCGGGAGCTGAACAAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGGGATCCGCCAC	960
Db	901	TTCCGGGAGCTGAACAAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGGGATCCGCCAC	960
Qy	961	CCCGCGGCTGAAGAAGAAAGAGACGTGACCGTGCTGGACGTGGGCGAGCGCTACTTC	1020
Db	961	CCCGCGGCTGAAGAAGAAAGAGACGTGACCGTGCTGGACGTGGGCGAGCGCTACTTC	1020
Qy	1021	AGCTGCCCCCTGGACGAGGACTTCCGCAAGTACACCGCTTTCACATCCCCAGCATCAAC	1080
Db	1021	AGCTGCCCCCTGGACGAGGACTTCCGCAAGTACACCGCTTTCACATCCCCAGCATCAAC	1080
Qy	1081	AAGGAGACCCCGGCATCGCTACCAAGTACACGTGCTGCCCCAGGGCTGGAAGGGCAGC	1140
Db	1081	AAGGAGACCCCGGCATCGCTACCAAGTACACGTGCTGCCCCAGGGCTGGAAGGGCAGC	1140
Qy	1141	CCGAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCGCAACCCC	1200
Db	1141	CCGAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCGCAACCCC	1200
Qy	1201	GAGATCGTGAATPACCA-----GGCCCCCTGTACGTGGGCGAGCATGTGGAGATCGGC	1260
Db	1201	GAGATCGTGAATPACCAAGTACATACGACGACCTGTACGTGGGCGAGCATGTGGAGATCGGC	1260
Qy	1261	CAGACCCGSCCAAGATCGAGGAGTGGCGAGCAGCTGCTGCCTGGGGCTTACCAACC	1320
Db	1261	CAGACCCGSCCAAGATCGAGGAGTGGCGAGCAGCTGCTGCCTGGGGCTTACCAACC	1320
Qy	1315	CCGCAAGAAGCACCAAGAGGAGCCCTTCTTCCTGCCCAT-----CGAGCTGCACCCC	1368
Db	1321	CCGCAAGAAGCACCAAGAGGAGCCCTTCTTCCTGTGGATGGGCTACGAGCTGCACCCC	1380
Qy	1369	GACAAGTGGACCTGGAGCCATCGAGCTGCCGAGAGAGAGAGCTGAGCCGTGAAGCAGC	1428
Db	1381	GACAAGTGGACCTGGAGCCATCGAGCTGCCGAGAGAGAGAGCTGAGCCGTGAAGCAGC	1440
Qy	1429	ATCCAGAAGCTGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCGGCATCAAGTG	1488
Db	1441	ATCCAGAAGCTGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCGGCATCAAGTG	1500
Qy	1489	CGCAGCTGTGCAAGCTGCTGGCGCGCGCAAGGCCCTGACCGACATCGTGCCCTGACC	1548
Db	1501	CGCAGCTGTGCAAGCTGCTGGCGCGCGCAAGGCCCTGACCGACATCGTGCCCTGACC	1560
Qy	1549	GAGAGGCCGAGCTGAGCTGGCGGAGAACCCGAGATCTCTGGGAGCCCGTGCACGGC	1608
Db	1561	GAGAGGCCGAGCTGAGCTGGCGGAGAACCCGAGATCTCTGGGAGCCCGTGCACGGC	1620

RESULT 4	AX427930	AX427930	9166 bp	DNA	linear	PAT 20-JUN-2002
LOCUS	AX427930					
DEFINITION	Sequence	168 from Patent	WO0232943.			
ACCESSION	AX427930					
VERSION	AX427930.1	GI:21538017				
KEYWORDS						
SOURCE		synthetic construct.				
ORGANISM		synthetic construct				
REFERENCE		artificial sequences.				
AUTHORS		1				
		Huang,Y. and Nabel,G.J.				

TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 168 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES Location/Qualifiers
source 1..9166
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2227 a 2662 c 2513 g 1764 t
ORIGIN

Query Match 81.4%; Score 2000.8; DB 6; Length 9166;
Best Local Similarity 91.3%; Pred. No. 1e-200;
Matches 2185; Conservative 0; Mismatches 182; Indels 25; Gaps 5;
QY 14 TGGCCGAGGCGATGAGCGAGCGCACACGCGCATCTGTGTCAGCGGAGCAACTTCA 73
DB 2961 TGGCCGAGGCGATGAGCAAGTGNACAAACACACATCATGTGACGCGGAGCAACTGCA 3020
QY 74 AGGCGCCCAAGCGCATCATCAAGTGTCTTCAACTGCGGCAAGGAGGCGCACATCGCCGCA 133
DB 3021 AGGCGCCCAAGCGCACCATCAAGTGTCTTCAACTGCGGCAAGGAGGCGCACATCGCCGCA 3080
QY 134 ACTGCGCGCGCCCGCCGCAAGAGGGCTGCTGGAGTCCGCGAAGGAGGCGCACAGATGA 193
DB 3081 ACTGCGCGCGCGCCCGCCGCAAGAGGGCTGCTGGAGTCCGCGAAGGAGGCGCACAGGTGA 3140
QY 194 AGGACTCACCGAGCGCGGCGCAACCTTCTCCGCGAGGACCTGGGCTTCCCGCAGGGCA 253
DB 3141 AGGACTCACCGAGCGACAGGTAA-TTTTTAGGGAAGATCTGGGCTTCCCAAGGGA 3199
QY 254 AGGCGCGCGAGTTCCCGAGCGAGAGCAACCGCGCCACACGCCCGCCAGCGGAGCTGC 313
DB 3200 AGGCGAGGAATTTCTTTCAGAGCAGACACAGAGCAACACGCCCGCCAGAGAGAGCTC 3259
QY 314 AGGTGCGCG- ----CGACAAACCCCGCAGAGGCGCGGCGGCGGCGGCGGCGGCGCA --- 364
DB 3260 AGGTGTTGGGAAGAGAGCAACCTTCCCTCTCAGAAGCAGGAGCGGATAGCAAGGAAGT 3319
QY 365 ---CCCTGAATCTCCCGCAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAGTGG 421
DB 3320 TATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCGACCCCTCGTCAACAATAAGATAG 3379
QY 422 GCGCGCAGATCAAGAGGCGCTGCTGGACACCGCGCGCGCGACACCGCTGCTGGAGAGA 481
DB 3380 GGGCGCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGCGACACCGCTGCTGGAGAGA 3439
QY 482 TGAGCCTGCGCGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTTCATCAAG 541
DB 3440 TGAACCTGCGCGCGCGCTGGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTTCATCAAG 3499
QY 542 TGGCGCAGTACGACCATCTGATCGAGATCTGCGGCAAGAGGCGCATCGGACCGCTGC 601
DB 3500 TGGCGCAGTACGACCATCTGATCGAGATCTGCGGCGCAAGGCGCATCGGACCGCTGC 3559
QY 602 TGATGCGCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCTGCA 661
DB 3560 TGGTGGCGCGCCACCGCGTGAACATCATCGCGCGCAACCTGCTGACCCAGATCGGCTGCA 3619
QY 662 CCCTGAACTTCCCGCATCAGCCCATCGAGACCGTGGCCGTGGAAGCTGAAGCCCGGATGG 721
DB 3620 CCCTGAACTTCCCGCATCAGCCCATCGAGACCGTGGCCGTGGAAGCTGAAGCCCGGATGG 3679
QY 722 ACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCGCTGACCGCCA 781
DB 3680 ACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCGCTGCTGGAGA 3739
QY 782 TCTGCGAGGAGATGGAAGAGGAGGCAAGATCATCAAGATCGGCGCGCGAGAACCCCTACA 841
DB 3740 TCTGCAACCGAGATGGAAGAGGAGGCAAGATCATCAAGATCGGCGCGCGAGAACCCCTACA 3799
QY 842 ACACCCCGTGTTCGCCCATCAAGAGAGGAGACACCAAGTGGCGGCAAGCTGGTGGACT 901

DB 3800 ACACCCCGTGTTCGCCCATCAAGAGAGGAGACACCAAGTGGCGCAAGCTGGTGGACT 3859
QY 902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGACGTGGGATCCCCCACC 961
DB 3860 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGACGTGGGATCCCCCACC 3919
QY 962 CCGCGCGCTGAAGAAGAAGACGCTGACCGTGTGAGCTGGGAGGAGCGCTTACTTCA 1021
DB 3920 CCGCGCGCTGAAGAAGAAGACGCTGACCGTGTGAGCTGGGAGGAGCGCTTACTTCA 3979
QY 1022 GCGTGCCCTGGAGGAGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCAACA 1081
DB 3980 GCGTGCCCTGGACAAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCAACA 4039
QY 1082 AGGAGACCCCGGATCGCTTACCAGTACACGCTGCTGCCCGAGGCTGGAGGCGAGCC 1141
DB 4040 ACGAGACCCCGGATCGCTTACCAGTACACGCTGCTGCCCGAGGCTGGAGGCGAGCC 4099
QY 1142 CCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCGCGCAACCCG 1201
DB 4100 CCGCATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCAAGCAGAACCCG 4159
QY 1202 AGATGCTGATCTACCA- ----GGCCCCCTGTAGCTGGGAGGACCTGGAGATCGGCC 1255
DB 4160 ACATGCTGATCTACCAAGTACATGGACCAACCTGTACGTGGCAGCGACTGGAGATCGGCC 4219
QY 1256 AGACCCCGCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGCTTTCACCAACC 1315
DB 4220 AGACCCCGCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGCTTTCACCAACC 4279
QY 1316 CCGACAAGAGCAGCAGAGGAGCGCCCTTCTTCCCTGCGCAT- ----CGAGCTGCACCCCG 1369
DB 4280 CCGACAAGAGCAGCAGAGGAGCGCCCTTCTTCTGTGATGGGTACGAGCTGCACCCCG 4339
QY 1370 ARAAGTGGACCTGCACCCCATCGAGCTGCCGAGAGAGGAGCTGACCGCTGAACGACA 1429
DB 4340 ACAAGTGGACCTGCACCCCATCGTGTGCCGAGAGAGGAGCTGACCGCTGAACGACA 4399
QY 1430 TCAGAAAGCTGGTGGGCAAGCTGAACCTGGCGCAGCAGATCTACCCGCGCATCAAGTGC 1489
DB 4400 TCAGAAAGCTGGTGGGCAAGCTGAACCTGGCGCAGCAGATCTACCCGCGCATCAAGTGC 4459
QY 1490 GCAGCTGTCAAGCTGCTCGCGCGCCCAAGGCCCTGACCGACATCTGCGCCCTGACCG 1549
DB 4460 GCAGCTGTCAAGCTGCTCGCGCGCCCAAGGCCCTGACCGAGGTGCTGCGCTGACCG 4519
QY 1550 AGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGGCGGAGCCCTGACACGGC 1609
DB 4520 AGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGAAGGAGCCCTGACACGGC 4579
QY 1610 TGTACTACGACCCCGCAGCAAGGAGCTGGTGGCGAGATCCAGAGCAGCGGCGCACCCAGT 1669
DB 4580 TGTACTACGACCCCGCAGCAAGGAGCTGATCGCGGAGATCCAGAGCAGCGGCGCACCGT 4639
QY 1670 GGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGA 1729
DB 4640 GGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCGCA 4699
QY 1730 TGGCAGCCCGCCACCAAGCAGCTGAAGCAGCTGACCGAGGCGGTGCAAGATGCCCA 1789
DB 4700 TGAAGGCGCGCCACCAAGCAGCTGAAGCAGCTGACCGAGGCGGTGCAAGATGCCCA 4759
QY 1790 TGGAGAGCATCTGTGGGCAAGACCCCAAGTTCGCCCTGCCCATCCAGAGGAGA 1849
DB 4760 CCGAGAGCATCTGTGGGCAAGACCCCAAGTTCAGCTGCCCATCCAGAGGAGA 4819
QY 1850 CTTGGGAGACCTGGTGGACCGACTACTTGGCAGGCGACCTGGATCCCCAGTGGAGTTCG 1909
DB 4820 CTTGGGAGGCTGGTGGACCGAGTACTTGGCAGGCGACCTGGATCCCCAGTGGAGTTCG 4879
QY 1910 TGAACACCCCGCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGGCCCATCATCTCGCG 1969

Db 4880 TGAACACCCCCCTTGTGAAGCTGTGGTACCGACTGGAGAGAGGCCCATCATCGGCG 4939

QY 1970 CCGAGACCTTCTACCTGGACGCGCGCAACCGCGAGACCAAGATCGCAAGCGCGCT 2029

Db 4940 CCGAGACCTTCTACCTGGACGCGCGCAACCGCGAGACCAAGCTGGCAAGCGCGCT 4999

QY 2030 AGTGACCGACCGGGCGCGGAGAGATCGTGAAGCTGACCGAGACCAACCAAGAGA 2089

Db 5000 AGTGACCGACCGGGCGCGGAGAGATCGTGAAGCTGACCGAGACCAACCAAGAGA 5059

QY 2090 CCGAGCTGCAGGCCATCGAGCTGGCCCTGCAGGACAGCGGCGAGAGTGAACATCGTGA 2149

Db 5060 CCGAGCTGCAGGCCATCGAGCTGGCCCTGCAGGACAGCGGCGAGAGTGAACATCGTGA 5119

QY 2150 CCGAGCGCAGTACCGCTTGGCCATCTCCAGCGCCAGCCGACAGAGCGAGCGAGC 2209

Db 5120 CCGAGCGCAGTACCGCTTGGCCATCTCCAGCGCCAGCCGACAGAGCGAGCGAGC 5179

QY 2210 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGTGC 2269

Db 5180 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGTGC 5239

QY 2270 CGCGCCCAAGGGCATCGCGGCAACGAGCAGATCGACAAGCTGTGAGCAAGGCGATCC 2329

Db 5240 CGCGCCCAAGGGCATCGCGGCAACGAGCAGGTGGAGCGGCTGTGAGCGCGCATCC 5299

QY 2330 GCAAGTGTCTTCTCGACGCGCATCGATGGCGGCATCGTGATCTACCAAGTA 2381

Db 5300 GCAAGTGTCTTCTCGACGCGCATCGACAAGGCCAGGAGGAGGACCAAGAA 5351

RESULT 5

AX427936

LOCUS AX427936 9788 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 174 from Patent WO232943.

ACCESSION AX427936

VERSION AX427936.1 GI:21538023

KEYWORDS synthetic construct.

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS Huang,Y. and Nabel,G.J.

TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization

JOURNAL Patent: WO 0232943-A 174 25-APR-2002;

GOVERNMENT OF THE UNITED STATES (US)

FEATURES

Location/Qualifiers

1..9788

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="plasmid pVR1012x/s containing HIV genes"

BASE COUNT 2377 a 2817 c 2695 g 1899 t

ORIGIN

Query Match 81.1%; Score 1992.4; DB 6; Length 9788;

Best Local Similarity 91.2%; Pred. No. 7.8e-200;

Matches 2181; Conservative 0; Mismatches 181; Indels 30; Gaps 5;

QY 14 TGGCCGAGGCGCATGAGCCAGGCGCACCAGCCACATCTCTGATGAGCGCAGCAACTTCA 73

Db 2961 TGGCCGAGGCGCATGAGCAAGGTGAACAACACCAATCATGATGAGCGCAGCAACTGCA 3020

QY 74 AGGGCCCCAAGCGCATCATCAAGTGTCTTCACTGGCGCAAGGAGGCGGCACATCGCCGCA 133

Db 3021 AGGGCCCCAAGCGCACCATCAAGTGTCTTCACTGGCGCAAGGAGGCGGCACATCGCCGCA 3080

QY 134 ACTCGCGCGCCCCCGCAAGAGGGCTGTCTGGAAGTGGCGCAAGGAGGCGGCACCAAGATGA 193

Db 3081 ACTCGCGCGCCCCCGCAAGAGGGCTGTCTGGAAGTGGCGCAAGGAGGCGGCACCAAGTGA 3140

QY 194 AGGACTGCACGCGCGCCAGGCGCAACTTCTTCCGCGAGGAGCTGGCTTCCCGCAGGCA 253

Db 3141 AGGACTGCACCGAGCGACAGGCTAA-----TAGGGAAGATCTGGCTTCCCAAGGGA 3194

QY 254 AGGCGCGCAGTTCCTCCACGAGCAGAAACCGCGCCCAAGAGCCCCACCGCGGAGCTGC 313

Db 3195 AGGCGCAGGAATTTCTTCAGAGCAGACAGAGCCCAACAGCCCCACCAAGAGAGACTTC 3254

QY 314 AGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGCGCGCCAGGCA--- 364

Db 3255 AGGTTTGGGAAGAGACACAACTCCCTCTCAGAGCAGAGCGCGGATAGACAAAGGAATG 3314

QY 365 ---CCCTGAACCTTCCCGCCAGATCAACCTGTGGAGCGCCCCCTGTGTGAGCATCAAGTGG 421

Db 3315 TATCCCTTTAGCTTCCTCAGATCACTCTTTGGCAGGACCCCTCTGCTACAATAAGATAG 3374

QY 422 GGGGCGAGATCAAGGAGGCCCTGCTGGACACCGCGCGCCGACACACCGTGTGGAGGAGA 481

Db 3375 GGGGCGAGCTGAAGGAGGCCCTTCTAGACACCGCGCGCCGACACACCGTGTGGAGGAGA 3434

QY 482 TGAGCTTGCCTCCCGCAAGTGGAAAGCCCAAGATGATCGCGGCATCGCGGCTTTCATCAAGG 541

Db 3435 TGAACCTGCCCGCGCGCTGGAAGCCCAAGATGATCGCGGCATCGCGGCTTTCATCAAGG 3494

QY 542 TCGCGCAGTACGACGAGATCTGTGATCGAGATCTGGCGCAAGAGGCCATCGCACCGTGC 601

Db 3495 TGGGCGAGTACGACGAGATCTGTGATCGAGATCTGGCGCCACAAGGCCATCGCACCGTGC 3554

QY 602 TGATCGGCGCCACCCCGCTGAACATCATCGCGCGCAACATGCTGACCGAGCTGGGCTGA 561

Db 3555 TGGTGGGCGCCACCCCGCTGAACATCATCGCGCGCAACCTGCTGACCGAGATCGGCTGA 3614

QY 662 CCCTGAACCTTCCCGCATCGACCCCATCGAGACCGTGCCTGAAGCTGAAGCCCGCGCATGG 721

Db 3615 CCCTGAACCTTCCCGCATCGACCCCATCGAGACCGTGCCTGAAGCTGAAGCCCGCGCATGG 3674

QY 722 ACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAATCAAGGCCCTTGACCGCA 781

Db 3675 ACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAATCAAGGCCCTTGTTGAGA 3734

QY 782 TCTGCGAGGAGATGAGAGGAGGCGCAAGATCACAAGATCGGCGCCCGAGACCCCTTACA 841

Db 3735 TCTGCGAGGAGATGAGAGGAGGCGCAAGATCAGGAAGATCGGCGCCCGAGACCCCTTACA 3794

QY 842 ACACCCCGCTGTTCGCCATCAAGAAGAGGACAGCACCAGTGGCGCAAGCTGGTGGACT 901

Db 3795 ACACCCCGCTGTTCGCCATCAAGAAGAGGACAGCACCAGTGGCGCAAGCTGGTGGACT 3854

QY 902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCAATCCCCCACC 961

Db 3855 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCAATCCCCCACC 3914

QY 962 CCGCGCGCTGAAGAAGAGAGCGCTGACCGTGTGGACGTGGCGGCGAGCGCTACTTCA 1021

Db 3915 CCGCGCGCTGAAGAAGAGAGCGCTGACCGTGTGGACGTGGCGGCGAGCGCTACTTCA 3974

QY 1022 GCGTCCCGCTGGAGGAGGACTTCCGCAAGTACACCGCTTACCATCCCCAGCATCAACA 1081

Db 3975 GCGTCCCGCTGGAGAGGACTTCCGCAAGTACACCGCTTACCATCCCCAGCATCAACA 4034

QY 1082 ACGAGACCCCGGCGATCGCTACCAAGTACAAGTGTCTCCCGAGGCGTGAAGGCGAGCC 1141

Db 4035 ACGAGACCCCGGCGATCGCTACCAAGTACAAGTGTCTCCCGAGGCGTGAAGGCGAGCC 4094

QY 1142 CCAGCATCTTCAGAGCAGCATGACCAAGATCTTGGAGCCCTTCCGCGCGCCCAACCCCG 1201

Db 4095 CCGCATCTTCAGAGTGCAGCATGACCAAGATCTTGGAGCCCTTCCGCGCAAGCAAGACCCCG 4154

QY 1202 AGATCGTGTATACCA-----GGCGCCCTGTAGCTGGGCGAGCGACTGGAGATCGGCC 1255

Db 4155 ACATCGTGTATACCAAGTACATGGACCACTGTAGTGGGCGAGCGACTGGAGATCGGCC 4214

QY 1256 AGCACCGCGCAAGATCGAGGAGTGCAGCAAGCACTGCTGGCTGGGCTTTCACACCC 1315

Db 4215 AGCACCGCAACCAAGATCGAGGAGTGCAGCAAGCACTGCTGGCTGGGCTTTCACACCC 4274


```
QY 662 CCCTGAACCTTCCCATATAGCCCATCGAGACCGTGCCTCCCTGAAGCTGAAGCCCGCATGG 721
Db 3623 CCCTGAACCTTCCCATATAGCCCATCGAGACCGTGCCTCCCTGAAGCTGAAGCCCGCATGG 3682
QY 722 ACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGAGAGAGATCAAGGCCCTGACCGCA 781
Db 3683 ACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGAGAGAGATCAAGGCCCTGATGGAGA 3742
QY 782 TCTGGAGAGATGAGAGAGGAGGCAAGATCACCAAGATGGCCCGCCGAGAACCCCTACA 841
Db 3743 TCTGCACCGAGATGAGAGAGGAGGCAAGATCAGCAAGATCGGCCCGCCGAGAACCCCTACA 3802
QY 842 ACACCCCGCTGTGCCATCAAGAAGAGAGACACCAAGTGGCGCAAGCTGTGGACT 901
Db 3803 ACACCCCGCTGTGCCATCAAGAAGAGAGACACCAAGTGGCGCAAGCTGTGGACT 3862
QY 902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGACGTGGGCAATCCCGCCACC 961
Db 3863 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGACGTGGGCAATCCCGCCACC 3922
QY 962 CCGCGCGCTGAAGAAGAGAGCGCTGACCGTGTGGAGTGGCGACGCTACTTCA 1021
Db 3923 CCGCGCGCTGAAGAAGAGAGCGCTGACCGTGTGGAGTGGCGACGCTACTTCA 3982
QY 1022 GCCTGCCCTGGACGAGGACTTCCGCAGTACACCGCTTCAACATCCCGACGATCAACA 1081
Db 3983 GCCTGCCCTGGACGAGGACTTCCGCAGTACACCGCTTCAACATCCCGACGATCAACA 4042
QY 1082 ACAGAGACCCCGGATCCGCTACCAAGTACAACGCTGCTGCCCGCCAGGCTGGAAGGCGAGCC 1141
Db 4043 ACAGAGACCCCGGATCCGCTACCAAGTACAACGCTGCTGCCCGCCAGGCTGGAAGGCGAGCC 4102
QY 1142 CCAGCATCTTCCAGAGAGCATGATACCAAGATTCCTGGAGCCCTTCCGGCCCGCCCAACCCCG 1201
Db 4103 CCAGCATCTTCCAGTGCAGCATGATACCAAGATTCCTGGAGCCCTTCCGGCAAGCAGAACCCCG 4162
QY 1202 AGATCGTATCTACCA-----GSCCCCTGTAGTGGGAGGACCTGGAGATCGGCC 1255
Db 4163 ACATCGTATCTACCAAGTATACCAAGTATACCAAGTATACCAAGTATACCAAGTATACCAAGT 4222
QY 1256 AGCACCAGCCAGATCGAGAGCTGGCGAAGACCTGCTGCTGGGCTTACCAACCC 1315
Db 4223 AGCACCAGCCAGATCGAGAGCTGGCGAAGACCTGCTGCTGGGCTTACCAACCC 4282
QY 1316 CCGACAGAGACCCAGAGAGGAGCCCTTCTGCCCCAT-----CGAGTGCACCCCG 1369
Db 4283 CCGACAGAGACCCAGAGAGGAGCCCTTCTGCTGGATGGGCTACGAGTGCACCCCG 4342
QY 1370 ARAAGTGCAGCTGCAGCCCATCAGCTGCCCGAGAGAGGAGTGGACCGTGAACGACA 1429
Db 4343 ARAAGTGCAGCTGCAGCCCATCAGCTGCCCGAGAGAGGAGTGGACCGTGAACGACA 4402
QY 1430 TCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCCGCATCAAGTGC 1489
Db 4403 TCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCCGCATCAAGTGC 4462
QY 1490 GCAGCTGTGCAAGCTGCTGCGGCGGCCAAGGCCCTGACACATCGTCCCTGACCG 1549
Db 4463 GCAGCTGTGCAAGCTGCTGCGGCGGCCAAGGCCCTGACACATCGTCCCTGACCG 4522
QY 1550 AGAGGCGGAGCTGGAGCTGGCGAGAACCCGAGATCTTGGCGGAGCCGCTGACCGGCG 1609
Db 4523 AGAGGCGGAGCTGGAGCTGGCGAGAACCCGAGATCTTGGCGGAGCCGCTGACCGGCG 4582
QY 1610 TGTACTACGACCCAGCAAGGAGCTGTGGCGCGAGATCCCAAGAGCAGGCGCACGACAGT 1669
Db 4583 TGTACTACGACCCAGCAAGGAGCTGTGGCGCGAGATCCCAAGAGCAGGCGCACGACAGT 4642
QY 1670 GGACCTACCAAGTCTACCAAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCAAGA 1729
Db 4643 GGACCTACCAAGTCTACCAAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCGCA 4702
QY 1730 TGGCAGCCGCCCCACACAGCAGCTGAAGAGCTGACCGAGGCGCTGCAGAGATCGCCA 1789
```

```
Db 4703 TGAAGGCGCCACACACACAGAGCTGAAGCAGCTGACCGAGCGCTGCAGAGATCGCCA 4762
QY 1790 TGGAGAGATCTGTATCTGGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAGGACA 1849
Db 4763 CCGAGAGATCTGTATCTGGGGCAAGACCCCAAGTTCAGCTGCCCATCCAGAAGGACA 4822
QY 1850 CCGTGGAGACCTGTGGACCGGACTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTGC 1909
Db 4823 CCGTGGAGACCTGTGGACCGGAGTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTGC 4882
QY 1910 TGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGCG 1969
Db 4883 TGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGCG 4942
QY 1970 CCGAGACCTTACCTGTGACGCGCGCCCAACCGGAGACCAAGATCGGCAAGCGCGCT 2029
Db 4943 CCGAGACCTTACCTGTGACGCGCGCCCAACCGGAGACCAAGATCGGCAAGCGCGCT 5002
QY 2030 ACGTGAACCGACCGGCGCGGCAAGAGATCGTGAGCTTGACCGAGACCAACCAAGACA 2089
Db 5003 ACGTGAACCGACCGGCGCGGCAAGAGATCGTGAGCTTGACCGAGACCAACCAAGACA 5062
QY 2090 CCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGAGTGAACATCGTGA 2149
Db 5063 CCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGTGGAGTGAACATCGTGA 5122
QY 2150 CCGAGACCGAGTACCGCTGGGCATCATCAGGCCCGCCAGCAAGAGCGAGAGCGAGC 2209
Db 5123 CCGAGACCGAGTACCGCTGGGCATCATCAGGCCCGCCAGCAAGAGCGAGAGCGAGC 5182
QY 2210 TGGTGAACCGAGATCGAGCAGCTGATCAAGAGAGAGAGTGTACCTGAGCTGGTGC 2269
Db 5183 TGGTGAACCGAGATCGAGCAGCTGATCAAGAGAGAGAGTGTACCTGAGCTGGTGC 5242
QY 2270 CCGGCCCAAGAGGATCGCGGCGACAGCAGATGACCAAGCTGTGAGCAAGGCGCATCC 2329
Db 5243 CCGGCCCAAGAGGATCGCGGCGACAGCAGAGTGTGAGCGGCTGTGAGCGCGCATCC 5302
QY 2330 GCAAGGTGCTTCTCTGGACCGCATCGATGGGCGCATCGTGTATCTACAGTA 2381
Db 5303 GCAAGGTGCTTCTCTGGACCGCATCGATGGGCGCATCGATGGGCGCATCGAGTA 5354
```

```
RESULT 7
AX427926 LOCUS AX427926 9194 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 164 from Patent WO0232943.
ACCESSION AX427926
VERSION AX427926.1 GI:21538013
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
JOURNAL genetic immunization
PATENT Patent: WO 0232943-A 164 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
LOCATION/Qualifiers
FEATURES
source 1. .9194
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2232 a 2672 c 2523 g 1767 t
ORIGIN
```

```
Query Match 80.4%; Score 1975; DB 6; Length 9194;
Best Local Similarity 90.9%; Pred. No. 5.2e-198;
Matches 2177; Conservative 0; Mismatches 190; Indels 28; Gaps 6;
QY 14 TGGCCGAGGCGCATGAGCCAGG---CCACCAGCGGCAACATCCTGATGCAGCGCAGCAACT 70
```

Db 2967 TGGCCGAGCCATGAGCGAGGTACCAACAGCGCCACCATCATGATGAGCGCGCAACT 3026
QY 71 TCAAGGGCCCCAAGCAGCATATCAAGTGTTCAACTTCGCGCAAGGAGGGCCACATCGCCC 130
Db 3027 TCGCAACACAGCGCAAGTCTGAAGTGTTCAACTTCGCGCAAGGAGGGCCACACCGCCC 3086
QY 131 GCAACTTGGCGCGCCCCCGCAAGAGGGTGTGGAAGTTCGCGCAAGGAGGGCCACACGA 190
Db 3087 GCAACTTGGCGCGCCCCCGCAAGAGGGTGTGGAAGTTCGCGCAAGGAGGGCCACACGA 3146
QY 191 TGAAGGACTGCACGAGCGCGCCAGCCAACTTCTTCGCGAGGAGACCTTGGCCCTTCCCCAGG 250
Db 3147 TGAAGGACTGCACGAGCGAGACGCTAA-TTTTTATGGAAGATCTGGCCCTTCCCCAAG 3205
QY 251 GCAAGGCCCGGAGTTCGCCAGCGAGGAGAGAACCGCGGCCAACACGCCACCAGCGCGAGC 310
Db 3206 GGAAGGCCAGGGAATTTCTTTCAGAGCAGACAGACCAACAGCCGCCACACGAAGAGAGC 3265
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGCGAGCGCAGGGCA 364
Db 3266 TTCAGGTTTGGGAAGAGAGACAACACTCCCTCTCAGNAGCAGAGCGCATAGACAGGAA 3325
QY 365 -----CCCTGAATTCGCCAGATCACCTGTGGCAGCGCCCTTGGTGAAGCATCAAGG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCGACCCCTCTGTCAACAATAAGA 3385
QY 419 TGGCGGCCAGATCAAGGAGGCCCTGTGTGACACCGCGCGCAGCAGACCGTGTGGAGG 478
Db 3386 TAGGGGCCAGCTGAAGGAGGGCCCTTCTACACACCGCGCGCAGCAGACCGTGTGGAGG 3445
QY 479 AGATGAGCCTGCCCGGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGGGCTTCATCA 538
Db 3446 AGATGAACTTGCCCGCGCTGGAAGCCCAAGATGATCGCGCGCATCGCGGGCTTCATCA 3505
QY 539 AGGTGCGCCAGTAGCAGCAGATCCTGTATCGAGATCTCGCGCAAGAGCGCATCGGCACCG 598
Db 3506 AGGTGCGCCAGTAGCAGCAGATCCTGTATCGAGATCTCGCGCCACAAAGGCCATCGGCACCG 3565
QY 599 TGTGATGCGCCACCCCGTGAACATCATCGSCGCCAATCATCTCACCACAGCTGGCT 658
Db 3566 TGTGTGTGGGCCCCCGTGAACATCATCGSCGCCAATCATCTCACCACAGATCGGCT 3625
QY 659 GCACCTTGAATTCGCCATCAGCCCATCGAGACCGTGCCTGTGAAGCTGAAGCCCGGCA 718
Db 3626 GCACCTTGAATTCGCCATCAGCCCATCGAGACCGTGCCTGTGAAGCTGAAGCCCGGCA 3685
QY 719 TGGAGGCCCGCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 3686 TGGAGGCCCGCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTTGTGG 3745
QY 779 CCATCTCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT 838
Db 3746 AGATCTGACGAGATGGAGNAGGAGGCAAGATCAGCAAGATCGGCCCGGAGAACCCCT 3805
QY 839 ACAACACCCCGTGTTCGCCATCAAGAAAGAGAGACACCAAGTGGCGCAAGTGTGG 898
Db 3806 ACAACACCCCGTGTTCGCCATCAAGAAAGAGAGACACCAAGTGGCGCAAGTGTGG 3865
QY 899 ACTTTCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGTGGGCATCCCC 958
Db 3866 ACTTTCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGTGGGCATCCCC 3925
QY 959 ACCCGCGGCTGAAGAAGAAGAGAGCTGACCGTGTGAGCTGGGCGACGCCCTACT 1018
Db 3926 ACCCGCGGCTGAAGAAGAAGAGAGCTGACCGTGTGAGCTGGGCGACGCCCTACT 3965
QY 1019 TCAGGTCGCCCTGGACGAGGACTTTCGCAAGTACACCGCCCTTCACCATCCCCAGCATCA 1078
Db 3986 TCAGGTCGCCCTGGACGAGGACTTTCGCAAGTACACCGCCCTTCACCATCCCCAGCATCA 4045
QY 1079 ACAAGGAGACCCCGGATCCGCTACCAAGTACACGCTGCTGCCCGCAGGGCTGGAAGGCA 1138

Db 4046 ACAACGAGACCCCGCATCCGCTACCAAGTACAACTGTGCTGCCAGGGGTGGAAGGGCA 4105
QY 1139 GCCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCGCAACC 1198
Db 4106 GCCCGCCATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCAAGCAGAAC 4165
QY 1199 CCAGAGTCTGTATCTACCA-----GGCCCCCTGTACGTGGGCGAGCAGCTGGAGATCG 1252
Db 4166 CCAGATCTGTATCTACCAAGTACATGACCACTGTACGTGGGCGAGCAGCTGGAGATCG 4225
QY 1253 GCAGACACCGCGCAGATGAGAGCTGCGCAAGCAGCTGCTCCGCTGGGGCTTCACCA 1312
Db 4226 GCAGACACCGCGCAGATGAGAGCTGCGCAAGCAGCTGCTCCGCTGGGGCTTCACCA 4285
QY 1313 CCCCCGACAAGAGCAGCAGAAAGGAGCCCTTCTCTGCCCAT-----CGAGTGCACC 1366
Db 4286 CCCCCGACAAGAGCAGCAGAAAGGAGCCCTTCTCTGTGTGATGGGCTACGAGTGCACC 4345
QY 1367 CCACAAAGTGGACCGTGCAGCCCATGAGCTGCCCCGAGAAGGAGAGCTGACCCGTGAACG 1426
Db 4346 CCACAAAGTGGACCGTGCAGCCCATGAGCTGCCCCGAGAAGGAGCAGCTGGACCGTGAACG 4405
QY 1427 ACATCAGAGAGCTGGTGGGCAAGCTGAACCTGGGCCAGCCAGATCTACCCGGCATCAAGG 1486
Db 4406 ACATCAGAGAGCTGGTGGGCAAGCTGAACCTGGGCCAGCCAGATCTACCCGGCATCAAGG 4465
QY 1487 TGGCCAGCTGTGCAAGCTGCTGCGCGCGCCAAAGGCCCTGACCGACATCGTGCCTCTGA 1546
Db 4466 TGGCCAGCTGTGCAAGCTGCTGCGCGCGCCAAAGGCCCTGACCGAGCTGCTGCCTCTGA 4525
QY 1547 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGCGAGATCTCTCGCGGAGCCCGTGACG 1606
Db 4526 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGCGAGATCTCTGAAGGAGCCCGTGACG 4585
QY 1607 GCGTGTACTAGCACCCCGCAGCAAGGAGCCTGTGTGGCCGAGATCTCAAGACAGGCGCCAGCACC 1666
Db 4586 GCGTGTACTAGCACCCCGCAGCAAGGAGCCTGTGTGGCCGAGATCTCAAGACAGGCGCCAGGCCC 4645
QY 1667 AGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGGAAGACCGGCAAGTACGCCA 1726
Db 4646 AGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGGAAGACCGGCAAGTACGCCA 4705
QY 1727 AGATGCGCACCGCCACACCAAGCAGCTGAGCAGCTGACCGAGCGCTGACCGAGCGCTGCGAGAGATCG 1786
Db 4706 GCATGGAAGGCGCCACCAACAGCAGCTGAAGCAGCTGACCGAGCGCTGCGAGAGATCG 4765
QY 1787 CCATGGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAAAG 1846
Db 4766 CCACCGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCAGCTGCCATCCAGAAAG 4825
QY 1847 AGACCTGGGAGACCTGTGGACCGACTACTTGGCAGGCCACCTGTGATCCCGAGTGGGAGT 1906
Db 4826 AGACCTGGGAGCCCTGTGGACCGAGTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT 4885
QY 1907 TCGTGAACACCCCGCCCTGTGTGAAGCTGTGTACCAAGCTGAGAGAGAGCCCATCATCG 1966
Db 4886 TCGTGAACACCCCGCCCTGTGTGAAGCTGTGTACCAAGCTGAGAGAGAGCCCATCATCG 4945
QY 1967 GCGCGGAGACCTTCTACGTGGACCGCGCCCAACCGCGAGACCAAGATCGCGCAAGGCGCG 2026
Db 4946 GCGCGGAGACCTTCTACGTGGACCGCGCCCAACCGCGAGACCAAGATCGCGCAAGGCGCG 5005
QY 2027 GCTACGTGACCGACCGGGCGCGGAGAGATCTGTGAGCCTGACCGAGACCAACACAGAG 2086
Db 5006 GCTACGTGACCGACCGCGCGCGGAGAGTGTGTGCCCTGACCGACACCAACACAGAG 5065
QY 2087 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCGAGGAGCGGAGCTGGAACATCG 2146
Db 5066 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCGAGGAGCGGCTGAGGTGACATCG 5125
QY 2147 TGACCGACGAGCAGTACCGCCCTGGGCATCTCCAGGCCCGCCAGCCAGAGCGAGAGCG 2206
Db 5126 TGACCGACGAGCAGTACCGCCCTGGGCATCTCCAGGCCCGCCAGCCAGAGCGAGAGCG 5185

QY 2207 AGCTGGTGAACAGATATCGAGCAGCTGATCAAGAGGAGAGAGTGTACCTGAGCTGGG 2266
|||||
Db 5186 AGCTGGTGAACAGATATCGAGCAGCTGATCAAGAGGAGAGAGTGTACCTGAGCTGGG 5245
|||||
QY 2267 TCCCGCGCCCAAGGGCAGCGCGGCAAGCAGCAGCAGATCGACAAGCTGGTGAGCAAGGCA 2326
|||||
Db 5246 TCCCGCGCCCAAGGGCAGCGCGGCAAGCAGCAGCAGCTGGTGAGCGCGCGCA 5305
|||||
QY 2327 TCCGCAAGGTGCTGTCTCTGAGCGCATCGATGGCGGATCGTGATCTACAGTA 2381
|||||
Db 5306 TCCGCAAGGTGCTGTCTCTGAGCGCATCGACAAGGCCAGGAGGACACAGAA 5360
|||||
RESULT 8
AX427925
LOCUS AX427925 9194 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 163 from Patent WO0232943.
ACCESSION AX427925
VERSION AX427925.1 GI:21538012
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
1 artificial sequences.
REFERENCE
AUTHORS Huang,Y. and Nabel,G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 163 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source Location/Qualifiers
1..9194
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2232 a 2671 c 2524 g 1767 t
ORIGIN
Query Match 80.3%; Score 1973.4; DB 6; Length 9194;
Best Local Similarity 90.9%; Pred. No. 7.6e-198;
Matches 2176; Conservative 0; Mismatches 191; Indels 28; Gaps 6;
QY 14 TGGCGGAGGCATGAGCCAGG---CCACGAGCGCCACACATCCTGATGAGCGCAGCAACT 70
|||||
Db 2967 TGGCGGAGGCATGAGCCAGGTTGACCAACAGCGCCACCATCATGATGAGCGCGCAACT 3026
|||||
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGTCTCAACTGGCGCAAGGAGGCCACATCGCCCC 130
|||||
Db 3027 TCCGCAACCGCGCAAGATCTGAAGTGTCTCAACTGGCGCAAGGAGGCCACACCGCCC 3086
|||||
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCCACCCAGA 190
|||||
Db 3087 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCCACCCAGA 3146
|||||
QY 191 TGAAGGACTGCACCGAGCGCCAGCGCAACTTCTTCCGCGAGGAGCTGGCCCTTCCCGCAGG 250
|||||
Db 3147 TGAAGGACTGCACCGAGCGCAGAGCTAA-TTTTATTAGGAGATCTGGCCCTTCCACAAAG 3205
|||||
QY 251 GCAAGGCGCGGAGTTCCTCCAGCGCAGAGCAACCGCCCAACAGCCCGCAGCGCGCAGC 310
|||||
Db 3206 GGAAGGCCAGGGAATTTCTTCAGAGCAGACCCAGAGCCCAAGCCCGCCACAGAGAGAGC 3265
|||||
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGGCGGCGGCGAGGCA 364
|||||
Db 3266 TTCAGGTTTGGGGAAGAGACAACAACCTCTCTCAGAGCAGGAGCGCGATAGACAAGAA 3325
|||||
QY 365 -----CCCTGAAGTTCCTCCAGATACCTGTGGCAGCGCCCTTGGTGAGCATCAAGG 418
|||||
Db 3326 CTGTATCTCTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCTGTACATAAAGA 3385
|||||
QY 419 TGGCGCGCCAGATCAAGAGGCGCTGTGGACACCGCGCGCGCAGCACACCTGTGTGAGG 478
|||||
Db 3386 TAGGGGCCAGCTGAAGAGGGCCCTTTAGACACCGCGCGCGCAGCACACCTGTGTGAGG 3445
|||||

QY 479 AGATGAGCTGTCGCCGCAAGTGAAGCCCAAGATGATCGCGGCATCGCGCGCTTCATCA 538
|||||
Db 3446 AGATGAACCTGTCGCCGCGCTGGAAGCCCAAGATGATCGCGGCATCGCGCGCTTCATCA 3505
|||||
QY 539 AGGTGCGCCAGTACGACCAAGATCCTGATGAGATCTGCGGCAAGAGGCGCATCGGCACCG 598
|||||
Db 3506 AGGTGGGCGAGTACGACCAAGATCCTGATCGAGATCTGCGGCGCACAGGCCATCGGCACCG 3565
|||||
QY 599 TGCTATCGGCGCCCGCCCGCTGACATCATCGGCGGCAACATGTCGACCCAGCTGGGCT 658
|||||
Db 3566 TGCTGTTGGGCGCCCGCCCGCTGAAATCATCGCGCGCAACCTGCTGACCCAGATCGGCT 3625
|||||
QY 659 GCACCTTGAACTTCCCGCATCGAGCCCATCGAGACCGTGCCTGAGCTGAAGCTGAAGCCCGCA 718
|||||
Db 3626 GCACCTTGAACTTCCCGCATCGAGCCCATCGAGACCGTGCCTGAGCTGAAGCCCGCA 3685
|||||
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGAGCCCTGACCG 778
|||||
Db 3686 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGAGCCCTGGTGG 3745
|||||
QY 779 CCATCTGCGAGGAGATGGAGAGGAGGCAAGATCACCAAGATCGGCGCCCGAGAACCCCT 838
|||||
Db 3746 AGATCTGCACCGAGATGGAGAGGAGGCAAGATCAGCAAGATCGGCGCCCGAGAACCCCT 3805
|||||
QY 839 ACAACACCCCGCTGTTCCGCATCAAGAGAGGAGCAGCACCAAGTGGCGCAAGCTGGTGG 898
|||||
Db 3806 ACAACACCCCGCTGTTCCGCATCAAGAGAGGAGCAGCACCAAGTGGCGCAAGCTGGTGG 3865
|||||
QY 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGAGTGCAGTGGCGCATCCCC 958
|||||
Db 3866 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGAGTGCAGTGGCGCATCCCC 3925
|||||
QY 959 ACCCGCGCGCTGAAGAGAGAGAGCGTGACCGTGTGACGAGTGGCGGAGCGCTACT 1018
|||||
Db 3926 ACCCGCGCGCTGAAGAGAGAGAGCGTGACCGTGTGACGAGTGGCGGAGCGCTACT 3985
|||||
QY 1019 TCAGCTGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 1078
|||||
Db 3986 TCAGCTGCCCTGGACAGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 4045
|||||
QY 1079 ACAAGGAGACCCCGGCATCCGCTACCAAGTACAAGTGTGCTGCCCGAGGCTGGAAGGCA 1138
|||||
Db 4046 ACAAGGAGACCCCGGCATCCGCTACCAAGTACAAGTGTGCTGCCCGAGGCTGGAAGGCA 4105
|||||
QY 1139 GCGCCAGCATCTTCAGAGAGCAGATGACCAAGATCCTTGAGCCCTTCGCGCGCCGCAACC 1198
|||||
Db 4106 GCGCCGCGCATCTTCAGTGCAGCATGACCAAGATCCTTGAGCCCTTCGCGCGCCGCAACC 4165
|||||
QY 1199 CCGAGATCGTGATCTACCA-----GGCGCCCTGTAGTGGCGAGCGACCTGGAGATCG 1252
|||||
Db 4166 CCGACATCGTGATCTACCAAGTACATGGAGACCTGTAGTGGCGAGCGACCTGGAGATCG 4225
|||||
QY 1253 GCCAGACCGCGCCCAAGATCGAGGAGTGCACAAGCAGCTGTGCTGCGCTGGGCTTCACCA 1312
|||||
Db 4226 GCCAGACCGCGCCCAAGATCGAGGAGTGCACAAGCAGCTGTGCTGCGCTGGGCTTCACCA 4285
|||||
QY 1313 CCGCGGACAAGACCCAGAGAGGCGCCCTTCTCTGCCAT-----CGAGCTGCACC 1366
|||||
Db 4286 CCGCGGACAAGACCCAGAGAGGCGCCCTTCTCTGGATGGGCTACGAGCTGCACC 4345
|||||
QY 1367 CCGACAAGTGACCGCTGCAGCCCATCGAGCTGCCGAGAGGAGCTGGACCGTGAACG 1426
|||||
Db 4346 CCGACAAGTGACCGCTGCAGCCCATCGTGTGCTGCCGAGAGGAGCTGGACCGTGAACG 4405
|||||
QY 1427 ACATCCAGAAAGCTGGTGGGCAAGCTGAACTGGGCGAGCAGATCTACCCCGCATCAAGG 1486
|||||
Db 4406 ACATCCAGAAAGCTGGTGGGCAAGCTGAACTGGGCGAGCAGATCTACCCCGCATCAAGG 4465
|||||
QY 1487 TGGCGCGCAGCTGTGCAAGCTGTGCGCGCGCCAGGCGCTTGACGAGATCTGTGCCCTTGA 1546
|||||
Db 4466 TGGCGCGCAGCTGTGCAAGCTGTGCGCGCGCCAGGCGCTTGACGAGCTGTGCCCTTGA 4525
|||||

QY 1547 CCAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGGAGATCTCTGCGGAGCCCGCTGCACG 1606
Db 4526 CCAGGAGGCCGAGCTGGAGCTGGCCGAGAACCGGAGATCTCTGCGGAGCCCGCTGCACG 4585
QY 1607 GCGTGACTACGACCCAGCAAGGACCTGGTGCCGAGATCCAGAGACGAGGCGCACGACC 1666
Db 4586 GCGTGACTACGACCCAGCAAGGACCTGGTGCCGAGATCCAGAGACGAGGCGCACGACC 4645
QY 1667 AGTGGACCTACGAGATCTACGAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1726
Db 4646 AGTGGACCTACGAGATCTACGAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCC 4705
QY 1727 AGATGGCAGCCGCCACCAACGAGCTGAAGCAGCTGACCGAGGCGCTGCAGAAAGATCG 1786
Db 4706 GCATGAAGGCGGCCACACACGAGCTGAAGCAGCTGACCGAGGCGCTGCAGAAAGATCG 4765
QY 1787 CCATGGAGAGCATCTGTGATCTGGGCAAGACCCCAAGTTCCGCTTGGCCCATCCAGAAAG 1846
Db 4766 CCACCGAGAGCATCTGTGATCTGGGCAAGACCCCAAGTTCCGCTTGGCCCATCCAGAAAG 4825
QY 1847 AGACCTGGGAGACCTGGTGACCGACTACTTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1906
Db 4826 AGACCTGGGAGGCTGGTGACCGAGTACTTGGCAGGCCACCTGGATCCCGAGTGGGAGT 4885
QY 1907 TCGTGAACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGGAGAGGAGCCCATCATCG 1966
Db 4886 TCGTGAACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGGAGAGGAGCCCATCATCG 4945
QY 1967 GCGCGAGACCTTCTAGTGGAGCGGCGCCCAACCGCGAGACCAAGATCGGCAAGGCCG 2026
Db 4946 GCGCGAGACCTTCTAGTGGAGCGGCGCCCAACCGCGAGACCAAGATCGGCAAGGCCG 5005
QY 2027 GCTAGCTGACCGAGCGGCGGCGCAGAAAGATCTGTAGCCCTGACCGAGACCAACACAGA 2086
Db 5006 GCTAGCTGACCGAGCGGCGGCGCAGAAAGTGTGCCCCCTGACCGACACCAACACAGA 5065
QY 2087 AGACCGAGCTGCAGGCGATCCAGCTGGCCCTGCAGGACAGCGGAGGAGGTGAACATCG 2146
Db 5066 AGACCGAGCTGCAGGCGATCCAGCTGGCCCTGCAGGACAGCGGCGCTGGAGGTGAACATCG 5125
QY 2147 TGACCGACACCGAGTACGCCCTGGGCATCATCCAGGCCACCGCCGACAGAGGAGCG 2206
Db 5126 TGACCGACACCGAGTACGCCCTGGGCATCATCCAGGCCACCGCCGACAGAGGAGGAGCG 5185
QY 2207 AGCTGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTTACTGTAGCTGGG 2266
Db 5186 AGCTGTGAGCCAGATCATCGACAGCTGATCAAGAGGAGAGGTTACTGTAGCTGGG 5245
QY 2267 TGCCCGCCCAAGGGGATCGGGGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGGCA 2326
Db 5246 TGCCCGCCCAAGGGGATCGGGGCAACGAGCAGGTTGGACGGCTGGTGAGCGCCGCA 5305
QY 2327 TCCGCAAGGTGCTTCTTGGACGGCATCGATGGCGGCATCTGTATCTACAGTA 2381
Db 5306 TCCGCAAGGTGCTTCTTGGACGGCATCGACAAAGGCCACGAGGAGGACACGAGAA 5360

RESULT 9
AX427927 LOCUS AX427927 12411 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 165 from Patent WO0232943.
ACCESSION AX427927
VERSION AX427927.1 GI:21538014
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 165 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)

FEATURES
source
Location/Qualifiers
1. .12411
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2938 a 3733 c 3476 g 2264 t
ORIGIN
Query Match 80.3%; Score 1973.4; DB 6; Length 12411;
Best Local Similarity 90.9%; Pred. No. 7.1e-198;
Matches 2176; Conservative 0; Mismatches 191; Indels 28; Gaps 6;
QY 14 TGGCCGAGGCCATGAGCCAGG---CCACCAGCGCCAACTCTCTGATGACGCGCAGCAACT 70
Db 2967 TGGCCGAGGCCATGAGCCAGGTTGACCAACAGCGCCACCATCATGATGACGCGGCAACT 3026
QY 71 TCAGAGGCCCCAAGCGGATCATCAAGTGTCTCAACTTCGGCAAGGAGGCGCACATCGCCC 130
Db 3027 TCGGCAACCAAGCGCAAGATCGTGAAGTGTCTCAACTTCGGCAAGGAGGCGCACACCCGCC 3086
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGTGGAGTGGCGCAAGGAGGGCCACCAGA 190
Db 3087 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGTGGAGTGGCGCAAGGAGGGCCACCAGA 3146
QY 191 TGAAGGACTCACCGAGCGCCAGCCAACTTCTTCCCGAGGACCTGGCCTTCCCCCAGG 250
Db 3147 TGAAGGACTCACCGAGCGACAGGCTAA-TTTTTTAGGGAAGATCTGGCCTTCCCAACAG 3205
QY 251 GCAAGGCCCGGAGTTCCCGAGCGGACAGAACCGCGCCCAACAGCCGCCACCGCGCCGAGC 310
Db 3206 GGAAGGCCAGGGAATTTTCTTCAGAGCAGACCAAGAGCCCAACAGCCGCCACCGAAGAGCG 3265
QY 311 TGCAGGTGCCCGG-----CGACAACCCCGCAGCGAGGCGCGCCGAGCGCCAGGCGCA 364
Db 3266 TTCAAGTTTGGGAGAGAGACAACACTCCCTCTCAGAGCAGGAGCGCGATAGACAAGAA 3325
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTCGGTGAGCATCAAGG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAACAAGA 3385
QY 419 TGGCGCGCCAGATCAAGGAGGCGCTGTGGACACCGCGCGGAGACACCTGTCTGGAGG 478
Db 3386 TAGGGGCGCAGCTGAAGAGGCGCCCTCTAGACACCGCGCGGCGACACCTGTCTGGAGG 3445
QY 479 AGATGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGATGGCGGCGATCGCGGCTTCATCA 538
Db 3446 AGATGAACTGCCCGCGCGCTGGAAAGCCCAAGATGATGGCGGCGATCGCGGCTTCATCA 3505
QY 539 AGGTGCCCGCAGTAGCAGCAGATCTGTGAGATCTGCGGCAAGAGCCATCGGCACCG 598
Db 3506 AGGTGGGCGCAGTAGCAGCAGATCTGTGAGATCTGCGGCAAGAGCCATCGGCACCG 3565
QY 599 TGTGTATCGCGCCCGCCCGTGAACATCATCGCGCCCAACATCTGACCCAGCTGGGCT 658
Db 3566 TGTGTGTGGGCGCCCGCCCGTGAACATCATCGCGCCCAACATCTGACCCAGATCGGCT 3625
QY 659 GCACCTGAACTTCCCGCAGTCCCGCAGACCGCTGCGCGTGAAGCTGAAGCCCGGCA 718
Db 3626 GCACCTGAACTTCCCGCAGTCCCGCAGACCGCTGCGCGTGAAGCTGAAGCCCGGCA 3685
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCGCCCTGACCGAGGAGAAATCAAGGCCCTGACC 778
Db 3686 TGGACGCGCCCAAGGTGAAGCAGTGGCGCCCTGACCGAGGAGAAATCAAGGCCCTGAGTGG 3745
QY 779 CCATCTCGGAGGAGATGAGAGGAGGAGGCAAGATCACCAAGATCGGCCCGCAAGCCCT 838
Db 3746 AGATCTGCACCGAGATGAGAGGAGGAGGCAAGATCACCAAGATCGGCCCGCAAGCCCT 3805
QY 839 ACAACACCGCCCGTTCGCCCATCAAGAGAGGAGACAGCACCAAGTGGCGCAAGCTGGTGG 898
Db 3806 ACAACACCGCCCGTTCGCCCATCAAGAGAGGAGACAGCACCAAGTGGCGCAAGCTGGTGG 3865
QY 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTCTTGGGAGGTGCAGCTGGGCATCCCC 958

Db 3866 ACTTCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGTGGGCATCCCC 3925
QY 959 ACCCGCGCGGCTGAAGAAGAAGAGCGTGAACGCTGCTGACGTGGGGCGACGCCCTACT 1018
Db 3926 ACCCGCGCGGCTGAACGAGAGAGAGCGTGAACGCTGCTGACGTGGGGCGACGCCCTACT 3985
QY 1019 TCAGCGTGGCCCTGGAGAGGACTTCCGCAAGTACACCGCCCTTACCATCCCCAGCATCA 1078
Db 3986 TCAGCGTGGCCCTGGACAAGACTTCCGCAAGTACACCGCCCTTACCATCCCCAGCATCA 4045
QY 1079 ACAAGAGACCCCGGATCCGCTACCAAGTACACAGTGTCTGCCAGGGCTGGAAGGCA 1138
Db 4046 ACAAGAGACCCCGGATCCGCTACCAAGTACACAGTGTCTGCCAGGGCTGGAAGGCA 4105
QY 1139 GCCCGAGCATTTCCAGAGCAGTATGACCAAGATCTCTGGAGCCCTTCCGCCCGCGCAACC 1198
Db 4106 GCCCGGCACTTCTCCAGTGACATGACCAAGATCTCTGGAGCCCTTCCGCCAGCAGAAC 4165
QY 1199 CCGAGATCGTGATCTACCA-----GGCCCCCTGTAGTGGGAGGACCTGAGATCG 1252
Db 4166 CCGACATCGTGATCTACCAAGTATGACCAAGTGTAGTGGGAGGACCTGAGATCG 4225
QY 1253 GCCAGCACCGCGCAAGTTCGAGAGTGGCGCAAGCACCCTGCTGGCTGGGGCTTCACCA 1312
Db 4226 GCCAGCACCGCGCAAGTTCGAGAGTGGCGCAAGCACCCTGCTGGCTGGGGCTTCACCA 4285
QY 1313 CCCCAGCAAGAACGACGAGGAGGCGCCCTTCTGCCCCAT-----CCAGCTGCACC 1366
Db 4286 CCCCAGCAAGAACGACGAGGAGGCGCCCTTCTGCCCCAT-----CCAGCTGCACC 4345
QY 1367 CCGACAAGTGACCGCTGCAGCCCATCGAGCTGCCGAGAGGAGCTGGACCGTGAACG 1426
Db 4346 CCGACAAGTGACCGCTGCAGCCCATCGAGCTGCCGAGAGGAGCTGGACCGTGAACG 4405
QY 1427 ACATCCAGAGCTGGTGGGCAAGCTGACCTGGGCGCAGCAGATCTACCCCGCATCAAG 1486
Db 4406 ACATCCAGAGCTGGTGGGCAAGCTGACCTGGGCGCAGCAGATCTACCCCGCATCAAG 4465
QY 1487 TGCCCGAGCTGTCAAGCTGTGCGCGCGCAAGGCGCTGACCGACATCTGCCCTGA 1546
Db 4466 TGCCCGAGCTGTCAAGCTGTGCGCGCGCAAGGCGCTGACCGAGGTGTGCCCTGA 4525
QY 1547 CCGAGAGGCGGAGCTGGAGCTGCGCGAGAACCGCGAGATCTCTCGGAGCGCCGTGCACG 1606
Db 4526 CCGAGAGGCGGAGCTGGAGCTGCGCGAGAACCGCGAGATCTCTGAAGGAGCGCGTGCACG 4585
QY 1607 GCGTGTACTAGCACCCAGCAAGACCTGTGTGCGCGAGATCCAGAACGAGGCGCACGAC 1666
Db 4586 GCGTGTACTAGCACCCAGCAAGACCTGTGTGCGCGAGATCCAGAACGAGGCGCACGAC 4645
QY 1667 AGTGGACCTACAGATCTACCGAGGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1726
Db 4646 AGTGGACCTACAGATCTACCGAGGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 4705
QY 1727 AGATGCGACCGCGCCACCAACGACGTGAAGAGCTGACCGAGGCGCTGCGAAGATCG 1786
Db 4706 GCATGAAGGCGCGCCACCAACGACGTGAAGAGCTGACCGAGGCGCTGCGAAGATCG 4765
QY 1787 CCATGGAGAGCATCGTGATCTGGGGCAAGACCCCGCAAGTTCGCGCTGCCATCCAGAGG 1846
Db 4766 CCACGAGAGCATCGTGATCTGGGGCAAGACCCCGCAAGTTCGCGCTGCCATCCAGAGG 4825
QY 1847 AGACTGGGAGACTGTGTGACCGACTACTTGGCAGGCGCACTGTGATCCCGAGTGGGAGT 1906
Db 4826 AGACTGGGAGCGCTGTGTGACCGAGTACTTGGCAGGCGCACTGTGATCCCGAGTGGGAGT 4885
QY 1907 TCGTGAACACCCCGCTGTGTGAAGCTGTGTACCAAGCTGGAGAAGGAGCGCCATCATCG 1966
Db 4886 TCGTGAACACCCCGCTGTGTGAAGCTGTGTACCAAGCTGGAGAAGGAGCGCCATCATCG 4945
QY 1967 GCGCGGAGACTTCTACGTGACGGCGCGCCAAACCGCGAGACCAAGATGGCGAAGGCCG 2026

Db 4946 GCGCGGAGACCTTCTACGTGACGCGCGCCCAACCCGAGACCAAGCTGGGCAAGGCCG 5005
QY 2027 GCTAGCTGACCGACCGCGGCGCGCAAGAAGATCGTGAAGCTGACCGAGACCAACACCGA 2086
Db 5006 GCTAGCTGACCGACCGCGGCGCGCAAGAAGTGGTGGCCCTGACCGACACCAACCGA 5065
QY 2087 AGACCGAGCTGCAAGGCCATTCAGCTGGCCCTGCGAGGACAGCGGCGAGGTGAACATCG 2146
Db 5066 AGACCGAGCTGCAAGGCCATTCAGCTGGCCCTGCGAGGACAGCGGCGTGGAGGTGAACATCG 5125
QY 2147 TGACCGAGACCGAGTACGCCCTGGGCATCATCCAGGCCCGACCCGACAGAGCGAGAGCG 2206
Db 5126 TGACCGAGACCGAGTACGCCCTGGGCATCATCCAGGCCCGACCCGACAGAGCGAGAGCG 5185
QY 2207 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGG 2266
Db 5186 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGG 5245
QY 2267 TGCCCGCGCCCAAGGGCATTCGCGGCAACGAGCAGATCGACAAGCTGTGTAGCAAGGCA 2326
Db 5246 TGCCCGCGCCCAAGGGCATTCGCGGCAACGAGCAGCTGTGGTGGAGCGCGCA 5305
QY 2327 TCCGCAAGGTGCTGTCTCTGACGCGCATCGATGGCGGCATCGTGATCTACCACTA 2381
Db 5306 TCCGCAAGGTGCTGTCTCTGACGCGCATCGACAAGGCCGAGGAGGACCGAGAA 5360

RESULT 10
AX427938
LOCUS AX427938 9785 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 176 from Patent WO0232943.
ACCESSION AX427938
VERSION AX427938.1 GI:21538025
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 176 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source
1. .9785
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2360 a 2843 c 2731 g 1851 t
ORIGIN

Query Match 80.2%; Score 1970; DB 6; Length 9785;
Best Local Similarity 90.9%; Pred. No. 1.7e-197;
Matches 2174; Conservative 0; Mismatches 185; Indels 33; Gaps 6;

QY 14 TGGCGGAGGCCATGAGCCAGGCGCACCGCGCAACATCTGTATGCGAGCGCAACTTCA 73
Db 2967 TGGCGGAGGCCATGAGCCAGGCTGCAGCAGCACCAACATCATGATGCGCGGCACTTCA 3026
QY 74 AGGCGCCCAACGCGATCATCAAGTGTCTTCAACTGCGGCAAGGAGGCGCACATCGCCCGCA 133
Db 3027 AGGCGCAAGCG---CATCAAGTGTCTCAACTGCGGCAAGGAGGCGCACCTTGGCCCGCA 3083
QY 134 ACTCGCGCGCCCGCGCAAGAGGCGCTGCTGGAAGTGGCGAAGGAGGCGCACAGATGA 193
Db 3084 ACTCGCGCGCGCTGCGCAAGAGGCGCTGCTGGAAGTGGCGAAGGAGGCGCACAGATGA 3143
QY 194 AGGACTGCACCGAGCGCCAGGCGCAACTTCTTCCGCGAGGAGCTGGCCCTTCCCGCAGGCA 253
Db 3144 AGGACTGCACCGAGCGGACAGGCTTA-----TAGGGAAGATCTGGCCCTTCCACAAAGGA 3197
QY 254 AGGCGCGGAGTTCCTCCAGCGAGCAGAAACCGCGCAACAGCGCCACCGAGCGCGAGCTGC 313

Db 3198 AGCCACGGGAATTTCTTCAGAGCAGACACACAGAGCCACACCCACAGAGAGGCTTC 3257
QY 314 AGTGGCGCG- ----CGACAACCCCGCAGCAGAGCGCGCGCGCAGCGCA- -- 364
Db 3258 AGTTTGGGAAGAGACAACACTCTCTCAGAGCAGAGCGCGATAGACAAGGAAGCTG 3317
QY 365 ---CCCTGAATTTCCCCAGATCACCTGTGGAGGCGCCCTGGTGGAGCATCAAGTGG 421
Db 3318 TATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAACAATAAGATAG 3377
QY 422 GCGCCAGATCAAGAGAGGCGCTGCTGGACACCGCGCGCAGACACCGTGTGGAGAGA 481
Db 3378 GGGCCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCAGCAGACCGTGTGGAGAGA 3437
QY 482 TGAGCTGCCCGCAAGTGAAGCCCAAGATGATCGCGGCATCGGCGGCTTCATCAAG 541
Db 3438 TGAACCTGCCCGCGCTGGAAGCCCAAGATGATCGCGGCATCGGCGGCTTCATCAAG 3497
QY 542 TGGCCAGTACGACACAGATCCTGATCGAGATCTGCGGCAAGAGGCGCATCGGCACCGTGC 601
Db 3498 TGGCCAGTACGACACAGATCCTGATCGAGATCTGCGGCCACAGGCCATCGGCACCGTGC 3557
QY 602 TGATGGCGCCACCCCGTGAACATCATGCGCGCGCAACATGCTGACCCAGCTGGGTGCA 661
Db 3558 TGGTGGCGCCACCCCGTGAACATCATGCGCGCGCAACCTGCTGACCCAGATCGGTGCA 3617
QY 662 CCCTGAATTTCCCATCAGCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCATGG 721
Db 3618 CCTGAATTTCCCATCAGCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCATGG 3677
QY 722 ACGGCCCAAGGTGAAGCAGTGCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCA 781
Db 3678 ACGGCCCAAGGTGAAGCAGTGCCTGACCGAGGAGAGATCAAGGCCCTGTTGGAGA 3737
QY 782 TCTGCGAGGAGTGAAGAGGAGGGCAAGTATCAACAGATCGCGCCCGGAGAACCCCTACA 841
Db 3738 TCTGCAACGAGATGAAGAGGAGGGCAAGATCAGCAAGATCGGCCCGGAGAACCCCTACA 3797
QY 842 ACACCCCGTGTTCGCATCAAGAAAGAGCAGACACCAAGTGGCGCAGCTGTGGACT 901
Db 3798 ACACCCCGTGTTCGCATCAAGAAAGAGGAGCAGACCAAGTGGCGCAGCTGTGGACT 3857
QY 902 TCGCGAGCTGAACAAAGCGCACCCAGGACTTCTGGAGGTGTCAGCTGGGCATCCCCACC 961
Db 3858 TCGCGAGCTGAACAAAGCGCACCCAGGACTTCTGGAGGTGTCAGCTGGGCATCCCCACC 3917
QY 962 CCGCCGCTGAAGAGAGAGAGCGTGNCCGTGTGAGCTGGGAGCGGCTACTTCA 1021
Db 3918 CCGCGGCTGAAGAGAGAGAGCGTGNCCGTGTGAGCTGGGAGCGGCTACTTCA 3977
QY 1022 GCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCTTCACCATCCCGCAGCATCAACA 1081
Db 3978 GCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCTTCACCATCCCGCAGCATCAACA 4037
QY 1082 ACAGACCCCGCGCATCCGCTACCACTACCAAGCTGTGCCCGCAGGCTGGAAGGGCAGCC 1141
Db 4038 ACAGACCCCGCGCATCCGCTACCACTACCAAGCTGTGCCCGCAGGCTGGAAGGGCAGCC 4097
QY 1142 CCAGCATCTCCAGAGCAGCATGACCAAGATCTCTGAGCGCTTCCGGCGCCGCAACCCCG 1201
Db 4098 CCGCCATCTCCAGTGCAGCATGACCAAGATCTCTGAGCGCTTCCCGCAGCAGAACCCCG 4157
QY 1202 AGATCGTGTACTACCA- ----GGCCCGCTGTACGTGGGCGCAGCTGGAGATCGGCC 1255
Db 4158 ACATCGTGTACTACAGTACATGACCACTGTACGTGGCAGCGACCTGGAGATCGGCC 4217
QY 1256 AGCACCGCGCAAGATCGAGGAGCTGCGCAAGCACTGTGCGCTGGGCTTCACCAACC 1315
Db 4218 AGCACCGCAAGATCGAGGAGCTGCGCAGCACTGTGCGCTGGGCTTCACCAACC 4277
QY 1316 CCGCAAGAGCAGCAGAGGAGGCGCCCTCTCTGCCCAT- ----CGAGCTGACCCCG 1369
Db 4278 CCGCAAGAGCAGCAGAGGAGGCGCCCTCTCTGCCCAT- ----CGAGCTGACCCCG 4337

QY 1370 ACAAGTGGACCTGACGCCCATCGAGCTGCCGAGAGGAGAGCTGACCGTGAACGACA 1429
Db 4338 ACAAGTGGACCTGACGCCCATCGTGTGCCGAGAGAGAGCTGACCGTGAACGACA 4397
QY 1430 TCCAGAGCTGTGGGCAAGCTGAACCTGGCCAGCCAGATCTACCCCGGATCAAGGTGC 1489
Db 4398 TCCAGAGCTGTGGGCAAGCTGAACCTGGCCAGCCAGATCTACCGCGGATCAAGGTGC 4457
QY 1490 GCCAGCTGTCAAGCTGTCTCGCGGCGCCAAAGGCCCTGACCGACATCGTGCCTGACGC 1549
Db 4458 GCCAGCTGTCAAGCTGTCTCGCGGCGCCAAAGGCCCTGACCGAGGTGTGCCCTGACGC 4517
QY 1550 AGGAGCGCGAGCTGGAGCTGGCCGAGAACCGCAGAGATCTCTGCGGAGCCGCTGCACCGCG 1609
Db 4518 AGGAGCGCGAGCTGGAGCTGGCCGAGAACCGCAGAGATCTCTGAAGAGCCCGCTGCACCGCG 4577
QY 1610 TGTACTACGACCCAGCAGAGCTGTGTGGCCGAGATCTCCAGAGCAGGCGCAGACCACT 1669
Db 4578 TGTACTACGACCCAGCAGAGCTGTATCGCCGAGATCTCCAGAGCAGGCGCAGGCGCAGT 4637
QY 1670 GGACCTACAGATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGA 1729
Db 4638 GGACCTACAGATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGA 4697
QY 1730 TGCACACCGCCACACACGACGTGAAGCAGCTGACCGAGGCGCTGCAGAAAGATCGCCA 1789
Db 4698 TGAAGGCGGCCACACACGACGTGAAGCAGCTGACCGAGGCGCTGCAGAAAGATCGCCA 4757
QY 1790 TGGAGAGCATCTGTATCTGGGCAAGACCCCAAGTTCCGCCCTGCCCATCCAGAAAGAGA 1849
Db 4758 CCGAGAGCATCTGTATCTGGGCAAGACCCCAAGTTCAAGCTGCCCATCCAGAAAGAGA 4817
QY 1850 CTTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACTTGATCCCGAGTGGGAGTTGC 1909
Db 4818 CTTGGGAGGCTGGTGGACCGACTACTGGCAGGCCACTTGATCCCGAGTGGGAGTTGC 4877
QY 1910 TGAACACCCCGCTGTGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGCGC 1969
Db 4878 TGAACACCCCGCTGTGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGCGC 4937
QY 1970 CCGAGACCTTCTACGTGGACCGCGCCCAACCGCAGACCAAGATCGGCAAGGCCGCT 2029
Db 4938 CCGAGACCTTCTACGTGGACCGCGCCCAACCGCAGACCAAGCTGGCAAGGCCGCT 4997
QY 2030 ACCTGACCGACCGGCGCGCAGAAAGTCTGTAGCCTGACCGAGACCAACCAACAGAGA 2089
Db 4998 ACCTGACCGACCGGCGCGCAGAAAGTGTGTGCCCTTGACCGACACCAACCAACAGAGA 5057
QY 2090 CCGAGCTGACGGCCATCCAGCTGGCCCTGCAGGACAGCGCAGCGAGGTGAACATCTGA 2149
Db 5058 CCGAGCTGACGGCCATCCAGCTGGCCCTGCAGGACAGCGGCCCTGGAGGTGAACATCTGA 5117
QY 2150 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCAGCCAGAGAGGAGCGAGCG 2209
Db 5118 CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCAGCCAGAGGAGCGAGCG 5177
QY 2210 TGGTGAACAGATCATCGACGCTGATCAAGAGAGAGAGGTGTACTGAGCTGGGTGC 2269
Db 5178 TGGTGAACAGATCATCGACGCTGATCAAGAGAGAGAGGTGTACTGAGCTGGGTGC 5237
QY 2270 CCGCCACAGGCGCATCGCGGCAAGCAGCAGATCGACAAGCTGGTGAAGAGGCGCATCC 2329
Db 5238 CCGCCACAGGCGCATCGCGGCAAGCAGCAGGTGGACGCCCTGGTGAAGGCGCGCATCC 5297
QY 2330 GCAAGTGTCTTCTTCCTGGACGCGATCGATGGCGGCGATCGTGTACTACAGATA 2381
Db 5298 GCAAGTGTCTTCTTCCTGGACGCGATCGACAAGGCCCGAGGAGGAGCAGAGAA 5349

RESULT 11

AX427933

LOCUS

9167 bp

DNA

linear

PAT 20-JUN-2002

DEFINITION	Sequence 171 from Patent WO0232943.
ACCESSION	AX427933
VERSION	AX427933.1 GI:21538020
KEYWORDS	
SOURCE	synthetic construct.
ORGANISM	artificial construct
REFERENCE	artificial sequences.
AUTHORS	1
TITLE	Huang, Y. and Nabel, G.J.
JOURNAL	Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
FEATURES	Patent: WO 0232943-A 171 25-APR-2002;
source	GOVERNMENT OF THE UNITED STATES (US)
BASE COUNT	Location/Qualifiers
ORIGIN	1..9167
	/organism="synthetic construct"
	/db_xref="taxon:32630"
	/notes="plasmid pVR1012x/s containing HIV genes"
	2225 a 2668 c 2517 g 1757 t
	Query Match 80.0%; Score 1966.6; DB 6; Length 9167;
	Best Local Similarity 90.7%; Pred. No. 3.9e-197;
	Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;
QY	14 TGGCCGAGGCCATAGCCAGG---CCACCAGCGCCACATCTCTGATGCAGCGCAGCAACT 70
Db	2964 TGGCCGAGGCCATAGCCAGGTTGACCAACAGCGCCACCATCATGATGAGCGCGCAACT 3023
QY	71 TCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGGCC 130
Db	3024 TCCGACACCGCGCAAGATCGTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACCGGCC 3083
QY	131 GCAACTCGCGGCCCCCGGCAAGAGGGCTGCTGGAAGTGGCGAAGGAGGGCCACCAGA 190
Db	3084 GCAACTCGCGGCCCCCGGCAAGAGGGCTGCTGGAAGTGGCGAAGGAGGGCCACCAGA 3143
QY	191 TGAAGGACTGCACGAGGCGCCAGGCCAACTCTTCGCGAGGAGACTGGCCCTTCCCCCAGG 250
Db	3144 TGAAGGACTGCACGAGGCGCAGAGGCTAA-----TAGGGAAGATCTGGCCCTTCCCACAAG 3197
QY	251 GCAAGGCGCGGAGTTGCCCGACGCGAGCAAGACCGGCCAACAGCCCCACAGCGCGCAGG 310
Db	3198 GGAAGGCGAGGATTTCTTCAGAGCAGACCCAGGCCAACAGCCCCACAGAGAGAGC 3257
QY	311 TGCAGGTGCGGG-----CGACAACCCCGCAGCGAGCGCGCGCGAGCGCCAGGGCA 364
Db	3258 TTCAGGTTTGGGAAGAGACAACACTCCCTCTCAGAAAGCAGGAGCGCATAGCAAGGAA 3317
QY	365 -----CCCTGAACCTTCCCAGATACCCCTGTGCGAGCGCCCCCTGGTGAGCATCAAG 418
Db	3318 CTGTATCCTTTAGTCTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAATAAGA 3377
QY	419 TGGCGGCCAGATCAAGAGGGCCCTGCTGGACACCGGCGCGCAGACACCCGTGTGGAGG 478
Db	3378 TAGGGGGCCAGCTGAAGAGGGCCCTTCTAGACACCGGCGCGCAGACACCGTGTGGAGG 3437
QY	479 AGATGACCTTCCCGGCAAGTGGAAAGCCCAAGATGATCGGGGGGATCGCGGGCTTCATCA 538
Db	3438 AGATGAACCTTCCCGGCGGCTTGAAGGCCAAGATGATCGGGGGCATCGCGGGCTTCATCA 3497
QY	539 AGGTGCGCCAGTAGCAGCCAGATCCTGATCAGATCTCGGCAAGAGGCCATCGGCACCG 598
Db	3498 AGGTGGGCCAGTAGCAGCCAGATCCTGATCGAGATCTCGGGGCCAAGGCCATCGGCACCG 3557
QY	599 TGTGTGATCGGCCACCCCGTGAACATCATCGGCCCGCAACATGCTGACCCAGCTGGGCT 658
Db	3558 TGTGTGTGGGCCACCCCGTGAACATCATCGGCCCGCAACCTGCTGACCCAGATCGGCT 3617
QY	659 GCACCTTGAACTTCCCATCAGCCCATCGAGACCGGTGCGCGTGAAGTGAAGCCCGGCA 718
Db	3618 GCACCTTGAACTTCCCATCAGCCCATCGAGACCGGTGCGCGTGAAGTGAAGCCCGGCA 3677


```
Db 4758 CCACCGAGACATCGTGATCTGGGGCAAGACCCCAAGTTCAAGCTGCCCATCCAGAAGG 4817
QY 1847 AGACTGGGAGACTGTGTGACCGACTTACTGGCAGGCCACTTGATCCCGAGTGGGAGT 1906
Db 4818 AGACTGGGAGGCTGTGTGACCGAGTACTTGGCAGGCCACTTGATCCCGAGTGGGAGT 4877
QY 1907 TCGTGAACACCCCCCTGTGTGAAGCTGTGGTACCAGCTGGAGAGAGCCCATCATCG 1966
Db 4878 TCGTGAACACCCCCCTGTGTGAAGCTGTGGTACCAGCTGGAGAGAGCCCATCATCG 4937
QY 1967 GCGCGGAGACCTTCTAGCTGGAGCGGCGGCCCAACCGCGAGACCAAGATCGGAAGCCG 2026
Db 4938 GCGCGGAGACCTTCTAGCTGGAGCGGCGGCCCAACCGCGAGACCAAGATCGGAAGCCG 4997
QY 2027 GCTACGTTGACCGACCGGGCGCGCAGAGATCGTGACCTTGACCGAGACCAACACAGA 2086
Db 4998 GCTACGTTGACCGACCGGGCGCGCAGAGTGTGGCCCTTGACCGAGACCAACACAGA 5057
QY 2087 AGACCGAGCTGACGAGCCATCCAGCTGGCCCTTGACGAGCAGCGGAGGTGAACATCG 2146
Db 5058 AGACCGAGCTGACGAGCCATCCAGCTGGCCCTTGACGAGCAGCGGCTGGAGGTGAACATCG 5117
QY 2147 TGACCGACAGCCAGTACGCTGGCCCTTGCCATCATCCAGGCCCGCAGCGAGGAGAGCG 2206
Db 5118 TGACCGACAGCCAGTACGCTGGCCCTTGCCATCATCCAGGCCCGCAGCGAGGAGAGCG 5177
QY 2207 AGCTGTGTGAACAGATCATCGACGAGCTGATCAAGAGGAGAGGTTACTGAGCTGGG 2266
Db 5178 AGCTGTGTGAACAGATCATCGACGAGCTGATCAAGAGGAGAGGTTACTGAGCTGGG 5237
QY 2267 TGCCCGCCCAACAGGCGATCGGGCGCAACGAGCAGATCGACAGCTGGTGAAGCAAGGCA 2326
Db 5238 TGCCCGCCCAACAGGCGATCGGGCGCAACGAGCAGGTTGACGGCTGGTGAAGCGGCA 5297
QY 2327 TCAGCAAGTGTCTTCTGTGGAGGATCATGATGGCGGATCGTGATCTACCAAGTA 2381
Db 5298 TCAGCAAGTGTCTTCTGTGGAGGATCATGATGGCGGATCGTGATCTACCAAGTA 5352

RESULT 12
AX427928
LOCUS AX427928 9170 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 166 from Patent W00232943.
ACCESSION AX427928
VERSION AX427928.1 GI:21538015
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G. J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 166 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
1..9170
/organism="synthetic construct"
/db_xref="taxon:32830"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 225 a 2669 c 2519 g 1757 t
ORIGIN

Query Match 80.0%; Score 1966.6; DB 6; Length 9170;
Best Local Similarity 90.7%; Pred. No. 3.9e-197;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

QY 14 TGGCCGAGGCGATGAGCAGG---CCACCGAGCGCCAACTCTGATCGACGCGAGCAACT 70
Db 2967 TGGCCGAGGCGATGAGCAGGTTACCCACAGCGCCACCATCATGATGACGCGGGCAACT 3026
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTCTTCAACTGCGGCAAGGAGGCGGCACATCGCCC 130
```

```
Db 3027 TCGCCAACCAAGCGCAAGATCGTGAAGTGTCTCAACTGCGCAAGGAGGGCCACACCGCCC 3086
QY 131 GCAACTGCCCGCGCCCGCCCAAGAAAGGCTGTCTGGAAGTGGCGCAAGGAGGGCCACACAGA 190
Db 3087 GCAACTGCCCGCGCCCGCCCAAGAAAGGCTGTCTGGAAGTGGCGCAAGGAGGGCCACACAGA 3146
QY 191 TGAAGGACTCACCGAGCGCGCAGGCCAACTTCTTCCCGAGAGACCTGGCTTCCCTCCCGCAGG 250
Db 3147 TGAAGGACTCACCGAGCGCGCAGGCCAACTTCTTCCCGAGAGATCTGGCTTCCACACAG 3200
QY 251 GCAAGGCGCGGAGTTCGCCAGCGAGAGAACCGCGCCAAACAGCCCAACAGCGCGCGAGC 310
Db 3201 GGAAGGCCAGGGAATTTCTTTCAGAGCAGACAGACCAACAGCCCAACAGCGCGCGAGAGAGC 3260
QY 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGGCGCGCGCGCGCGAGCGCGCAGGCA 364
Db 3261 TTCAGGTTTGGGGAAGAGACAACTCCCTCTCAGAGCAGGAGCGGATAGACAAAGGAA 3320
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTGTGTGAGCATCAAGG 418
Db 3321 CTGTATCTTTAGCTTCCCTCAGATCACTTCTTGGCAGCGACCCCTGTGTCAATATAGA 3380
QY 419 TGGCGGCCAGATCAAGGAGGCGCTGTGTGACACCGCGCGCGCGAGCACACCGTGTGAGG 478
Db 3381 TAGGGGCCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGAGCACACCGTGTGAGG 3440
QY 479 AGATGAGCTTCCCGGCAAGTGGAGCCCAAGATGATGCGGGCGCATCGGGCTTTCATCA 538
Db 3441 AGATGAACTTCCCGGCGCGCTTGGAAAGCCCAAGATGATGCGGGCGCATCGGGCTTTCATCA 3500
QY 539 AGGTGCGCCAGTACGACGAGATCTGTGATCGAGATCTCGGCGCAAGAGCGCATCGGCACCG 598
Db 3501 AGGTGCGCCAGTACGACGAGATCTGTGATCGAGATCTCGGCGCAAGAGCGCATCGGCACCG 3560
QY 599 TGTGATGCGCGCCACCGCGTGAACATCATCGGCGCGCAACATGCTGACCGAGCTGGGCT 658
Db 3561 TGTGTTGGCGCCCGCCCGCTGAAACATCATCGGCGCGCAACCTGCTGACCGAGATCGGCT 3620
QY 659 GCACCTGAACTTCCCGCATCGACCCCATCGAGCCGTGCGCGTGAAGCTGAAGCGCGCGCA 718
Db 3621 GCACCTGAACTTCCCGCATCGACCCCATCGAGCCGTGCGCGTGAAGCTGAAGCGCGCGCA 3680
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 3681 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGTTGG 3740
QY 779 CCATCTCGAGGAGATGGAGAGGCGCAAGATTCACCAAGATCGGCCCGCGAGAACCCCT 838
Db 3741 AGATCTGCACCGAGATGGAGAGGCGCAAGATCAGCAAGATCGGCCCGCGAGAACCCCT 3800
QY 839 ACACACCGCCCGTTCGCCCATCAAGAGAGGAGACACCAAGTGGCGCAAGCTGGTGG 898
Db 3801 ACAACACCGCCCGTTCGCCCATCAAGAGAGGAGACACCAAGTGGCGCAAGCTGGTGG 3860
QY 899 ACTTCCCGGAGCTCAACAAGCGACCCAGGACTTCTTGGAGTGCAGTGGGCTATCCCGC 958
Db 3861 ACTTCCCGGAGCTGAACAAGCGACCCAGGACTTCTTGGAGTGCAGTGGGCTATCCCGC 3920
QY 959 ACCCGCGCGCTGAAGAAGAGAGCGGTGACCGCTGTGTGGAGTGGGCGACCGCTACT 1018
Db 3921 ACCCGCGCGCTGAAGAAGAGAGCGGTGACCGCTGTGTGGAGTGGGCGACCGCTACT 3980
QY 1019 TCAGCTGCGCCCTGGAGGAGGACTTCCGCAAGTACACCGCTTCACCATCCCCAGCATCA 1078
Db 3981 TCAGCTGCGCCCTGGAGAGGAGGACTTCCGCAAGTACACCGCTTCACCATCCCCAGCATCA 4040
QY 1079 ACAACGAGACCGCGGCTCCGCTACCAAGTACACGCTGCTGCCCGAGGCTGGAGAGGCA 1138
Db 4041 ACACGAGACCGCGGCTCCGCTACCAAGTACACGCTGCTGCCCGAGGCTGGAGAGGCA 4100
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCTGTGGAGCGCTTCCGCGCGCGCAACC 1198
```


Db 4101 GCGCGCCATCTTCCAGTGCAGCATGACCAAGATCCTTGAGCCCTTCGCGAAGCAGAAC 4160
Qy 1199 CCGAGATCTGATCTACCA-----GGCCCCCTCTAGTGGGACGACCTGGAGATCG 1252
Db 4161 CCGACATCTGATCTACCAAGTACATGACACCTGTAGTGGGACGACCTGGAGATCG 4220
Qy 1253 GCGAGCACCGCGGCAAGATCGAGGAGCTGCGCAAGCACCTCTGCGCTGGGGCTTCAACA 1312
Db 4221 GCGAGCACCGCGGCAAGATCGAGGAGCTGCGCAAGCACCTCTGCGCTGGGGCTTCAACA 4280
Qy 1313 CCGCGCAAGAACGACAGAGAGCGCCCTTCTCTCCCAT-----CGAGCTGCACC 1366
Db 4281 CCGCGCAAGAACGACAGAGAGCGCCCTTCTCTGATGGGCTTACGAGCTGCACC 4340
Qy 1367 CCGACAAGTGCAGCGTGCAGCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG 1426
Db 4341 CCGACAAGTGCAGCGTGCAGCCATCGTGTCTGCCGAGAGAGAGCTGGACCGTGAACG 4400
Qy 1427 ACATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGAGCCAGATCTACCCCGCATCAAG 1486
Db 4401 ACATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGAGCCAGATCTACCGCGCATCAAG 4460
Qy 1487 TGCGCCAGCTGTCAAGCTGTCTGCGCGCGCAAGCGCTTACCGACACATCTGCGCCCTGA 1546
Db 4461 TGCGCCAGCTGTCAAGCTGTCTGCGCGCGCAAGCGCTTACCGAGGCTGTGCGCCCTGA 4520
Qy 1547 CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCCTGCGGAGCGCGTGACG 1606
Db 4521 CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCCTGAGGAGCGCGTGACG 4580
Qy 1607 GCGTGTACTAGCACCCAGCAAGCAAGCACTGTGTGGCCGAGATCCAGAAAGCGCCACGACC 1666
Db 4581 GCGTGTACTAGCACCCAGCAAGCAAGCACTGTATCGCGGAGATCCAGAAAGCGCGGCGCC 4640
Qy 1667 AGTGAGCTTACAGATCTACCAAGAGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1726
Db 4641 AGTGAGCTTACCAAGATCTACCAAGAGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCC 4700
Qy 1727 AGATGCGCAACCGCCACCAACACGCTGAAGCAGCTGACCGAGGCGGTGCGAGAGATCG 1786
Db 4701 GCATGAAGGGCGCCACACCAACGACGTGAAGCAGCTGACCGAGGCGGTGCGAGAGATCG 4760
Qy 1787 CCATGGAGAGCATGCTGATCTGGGGCAGAGACCCCAAGTTCCGGCTGCGCCATCCAGAAG 1846
Db 4761 CCACGAGAGCATGCTGATCTGGGGCAGAGACCCCAAGTTCAAGCTGCCATCCAGAAGG 4820
Qy 1847 AGACCTGGGAGACCTGCTGAGCAGCTTACTGGCAGGCGCACTGGATCCCGAGTGGGAGT 1906
Db 4821 AGACCTGGGAGGCTGCTGAGCAGCTTACTGGCAGGCGCACTGGATCCCGAGTGGGAGT 4880
Qy 1907 TCGTGAACACCCCGCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGGCCATCATCG 1966
Db 4881 TCGTGAACACCCCGCCCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGGCCATCATCG 4940
Qy 1967 GCGCGGAGACCTTCTAGCTGGAGCGCGCGGCAACCGCGAGACCAAGATCGGCAAGGCG 2026
Db 4941 GCGCGGAGACCTTCTAGCTGGAGCGCGCGGCAACCGCGAGACCAAGATCGGCAAGGCG 5000
Qy 2027 GCTACGTGACCGAGCGCGCGGCGGAGAGATCGTGAAGCTGAGTGAAGAGAGGAGTGAACATCG 2086
Db 5001 GCTACGTGACCGAGCGCGCGGCGGAGAGGTTGGTGGCCCTGAGCGAGACCAACAGAG 5060
Qy 2087 AGACCGAGCTGCAGGCGCATCAGCTGGCCCTTGCAGGACAGCGGAGGAGTGAACATCG 2146
Db 5061 AGACCGAGCTGCAGGCGCATCAGCTGGCCCTTGCAGGACAGCGGCGTGGAGTGAACATCG 5120
Qy 2147 TGACCGACAGCAGTACGCGCTGGGCGATCATTCAGGCGCCAGCGCGACAGAGGAGAGCG 2206
Db 5121 TGACCGACAGCAGTACGCGCTGGGCGATCATTCAGGCGCCAGCGCGACAGAGGAGAGCG 5180
Qy 2207 AGCTGGTGAACACAGATCATCGACAGCTGATCAAGAGGAGAGAGTGTACTGAGCTGGG 2266
Db 5181 AGCTGGTGAACACAGATCATCGACAGCTGATCAAGAGGAGAGAGTGTACTGAGCTGGG 5240

Qy 2267 TCCCGCCACAAAGGCGATCGCGCGAAGCAGACAGATCGACAAGCTGGTGACCAAGGCA 2326
Db 5241 TCCCGCCACAAAGGCGATCGCGCGAAGCAGACAGATCGACAAGCTGGTGACCGCGCA 5300
Qy 2327 TCCGCAAGTGTCTTCTCTGACGCGCATGATGCGGCGATCTGATCTACCAAGTA 2381
Db 5301 TCCGCAAGTGTCTTCTCTGACGCGCATGACGCAAGGCCAGGAGGACGACGAGAA 5355

RESULT 13
AX427935
LOCUS AX427935 9782 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 173 from Patent WO0232943.
ACCESSION AX427935
VERSION AX427935.1 GI:21538022
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Huang,Y. and Nabel,G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 173 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source Location/Qualifiers
1..9782
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2339 a 2895 c 2719 g 1829 t
ORIGIN

Query Match 80.0%; Score 1966.6; DB 6; Length 9782;
Best Local Similarity 90.7%; Pred. No. 3.9e-197;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

Qy 14 TGGCGGAGGCGCATGAGCCAGG---CCACCAGGCGCAACATCTGTATGATGAGCGCAGCACT 70
Db 3579 TGGCGGAGGCGCATGAGCCAGGCGTGAAGTCTTCAACTGGCGCAAGGAGGCGCCACACT 3638
Qy 71 TCAAGGCGCCCAAGCGCATCATCAAGTCTTCAACTGGCGCAAGGAGGCGCCACATCGGCC 130
Db 3639 TCCGCAACACGCGCAAGATCTGAAGTCTTCAACTGGCGCAAGGAGGCGCCACACTGCC 3698
Qy 131 GCAACTGCGCGCGCCCGCGCAAGAGGCGTCTGGAAGTGGCGCAAGAGGCGCCACACAGA 190
Db 3699 GCAACTGCGCGCGCCCGCGCAAGAGGCGTCTGGAAGTGGCGCAAGAGGCGCCACACAGA 3758
Qy 191 TGAAGGACTGCACGAGCGCCAGCGCAACTTCTTCCGCGAGGAGCTGCGCTTCCCGCAGG 250
Db 3759 TGAAGGACTGCACGAGCGCGCCAGCGCAACTTCTTCCGCGAGGAGCTGCGCTTCCCGCAGG 3812
Qy 251 GCAAGCGCGCGAGTTTCCCGAGCGAGAGAACCGCGCAACAGCGCCACACAGCGCGCAGC 310
Db 3813 GGAAGCGCGAGGATTTCTTCAGAGCAGACAGCAGCCACAGCCCGCCACACAGAGAGAGC 3872
Qy 311 TGCAGTGTGCGGG-----CGACAACCGCCGAGCGCGCGCGCGCGCGCGCGCGCGCA 364
Db 3873 TTCAGTGTGCGGGAGAGACAACTCCCTCTCAGAAGCAGGAGCGCCATAGACAAGGAA 3932
Qy 365 -----CCCTGAACCTTCCCGCAGATCACCTTGTGGCAGCGCCCTTGTGAGCATCAAG 418
Db 3933 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTTGGCAGCGACCCCTCGTCAATAAAGA 3992
Qy 419 TGGCGCGCGAGATCAAGAGGAGCGCTTGTGGACACCGCGCGCGCGCGCGCGCGCGCGCG 478
Db 3993 TAGGGCGCGAGCTGAAGAGCGCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 4052
Qy 479 AGATGAGCTGCCCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCAATCA 538
Db 4053 AGATGAGCTGCCCGCGCTTGGAAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCAATCA 4112

QY 539 AGGTGGCCAGTACGACCATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
DB 4113 AGGTGGCCAGTACGACCATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 4172
QY 599 TGTGTATGCGCCACCCCGTGAACATCATCGGCGCAACATCTGACCCAGCTGGGCT 658
DB 4173 TGTGTGCGCCACCCCGTGAACATCATCGGCGCAACATCTGACCCAGCTGGGCT 4232
QY 659 GCACCTGAATCTCCCATCAGCCCATCGAGACCGTGCCTGAAAGTGAAGCCCGGCA 718
DB 4233 GCACCTGAATCTCCCATCAGCCCATCGAGACCGTGCCTGAAAGTGAAGCCCGGCA 4292
QY 719 TGAAGCCCGCAAGTGAAGCATCTGAGCATCTGAGCGCAACCTGCTGACCCAGCTGACCG 778
DB 4293 TGAAGCCCGCAAGTGAAGCATCTGAGCATCTGAGCGCAACCTGCTGAGCGCAACCTGCTG 4352
QY 779 CCATCTCGAGAGAGTGAAGAGGAGGCAAGATCAACCAAGATCGGCCCGGCAAGCCCT 838
DB 4353 AGATCTGCACCGAGATGGAAGAGGAGGCAAGATCAACCAAGATCGGCCCGGCAAGCCCT 4412
QY 839 ACAACACCCCGTGTGCGCATCAAGAGAGGAGGCAAGATCGGCCCGGCAAGCTGGTG 898
DB 4413 ACAACACCCCGTGTGCGCATCAAGAGAGGAGGCAAGATCGGCCCGGCAAGCTGGTG 4472
QY 899 ACTTCGCGAGCTGAACAGCGCACCCAGCATCTTCTGGAGGTGCAGCTGGGCATCCCC 958
DB 4473 ACTTCGCGAGCTGAACAGCGCACCCAGCATCTTCTGGAGGTGCAGCTGGGCATCCCC 4532
QY 959 ACCCGCGCGCTGAAGAGAGAGAGCGTGCAGCTGCGTGGAGCGGCGAGCGCTACT 1018
DB 4533 ACCCGCGCGCTGAAGAGAGAGAGCGTGCAGCTGCGTGGAGCGGCGAGCGCTACT 4592
QY 1019 TCAGCTGCCCCCTGGAGAGGACTTCGCAAGTACACCGCTTACCCATCCCCAGCATCA 1078
DB 4593 TCAGCTGCCCCCTGGAGAGGACTTCGCAAGTACACCGCTTACCCATCCCCAGCATCA 4652
QY 1079 ACAACAGAGCCCGGATCCGCTACCAAGTACACCGCTTACCCATCCCCAGGCA 1138
DB 4653 ACAACAGAGCCCGGATCCGCTACCAAGTACACCGCTTACCCATCCCCAGGCA 4712
QY 1139 GCGCCAGCATCTTCAGAGAGCATGACCAAGATCTCGAGCGCTTCGCGCGCGCAAC 1198
DB 4713 GCGCCAGCATCTTCAGAGAGCATGACCAAGATCTCGAGCGCTTCGCGCGCGCAAC 4772
QY 1199 CCGAGATCGTATCTACCA -----GCGCCCGCTGACGTGGCAGCGCTGGAGATCG 1252
DB 4773 CCGAGATCGTATCTACCAAGTACATGGACCAACCTGTACGTGGGCGAGCGCTGGAGATCG 4832
QY 1253 GCCAGCACCGCGCAAGATCGAGAGCTGGCGCAAGCACCTGCTGCGTGGGCGCTTACCA 1312
DB 4833 GCCAGCACCGCGCAAGATCGAGAGCTGGCGCAAGCACCTGCTGCGTGGGCGCTTACCA 4892
QY 1313 CCGCCGACAGAGACACAGAGAGCGCGCGCTTCCTGCGCAT -----CGAGCTGAC 1366
DB 4893 CCGCCGACAGAGACACAGAGAGCGCGCGCTTCCTGCGCATGGCTAGAGCTGACCG 4952
QY 1367 CCGACAGTGGACCGTGCAGCCATCGAGCTGCGCGAGAGGAGAGCTGGACCGTGAACG 1426
DB 4953 CCGACAGTGGACCGTGCAGCCATCGTGTGCGCGAGAGGAGAGCTGGACCGTGAACG 5012
QY 1427 ACATCCAGAGAGCTGGTGGCAAGTGAACCTGGGCGACCGCATCTACCCCGCATCAAG 1486
DB 5013 ACATCCAGAGAGCTGGTGGCAAGTGAACCTGGGCGACCGCATCTACCCCGCATCAAG 5072
QY 1487 TGCAGAGCTGTGCAAGCTGTGCGCGCGCGCAAGCGCTGACCGCATCTGCGCCCTGA 1546
DB 5073 TGCAGAGCTGTGCAAGCTGTGCGCGCGCGCAAGCGCTGACCGCATCTGCGCCCTGA 5132
QY 1547 CCGAGAGGCGGAGCTGAGCTGCGCGAGAGCGCGAGATCTCTGCGGAGCGCGCTGACG 1606
DB 5133 CCGAGAGGCGGAGCTGAGCTGCGCGAGAGCGCGAGATCTCTGAGAGCGCGCGCTGACG 5192

QY 1607 GCCTGTACTACGACCCCAAGGACCTGCTGGCCGAGATCCAGAGCGGCGCACGACC 1666
DB 5193 GCCTGTACTACGACCCCAAGGACCTGATCGCGAGATCCAGAGCGGCGCACGACC 5252
QY 1667 AGTGGACCTTACCAGATCTTACCAGAGCGCTTCAAGAACCTTGAAGACCGGAAGTACGCCA 1726
DB 5253 AGTGGACCTTACCAGATCTTACCAGAGCGCTTCAAGAACCTTGAAGACCGGAAGTACGCCA 5312
QY 1727 AGATGGCACCCGCCACACAGCATGTAAGCAGCTGACCGAGGCGCTGCAGAAAGATCG 1786
DB 5313 GCATGAAGGCGGCCACACCAAGCAGCTGAAGCAGCTGACCGAGGCGCTGCAGAAAGATCG 5372
QY 1787 CCATGGAGACATCTGATCTGGGGCAAGACCCCAAGTTCCGCTTCCCATCCAGAAAG 1846
DB 5373 CCACCGAGACATCTGATCTGGGGCAAGACCCCAAGTTCAAGCTGCCCATCCAGAAAG 5432
QY 1847 AGACCTGGGAGACCTTGTGTGACCGACTTACTGGCAGGCGACCTTGGATCCCGAGTGGAGT 1906
DB 5433 AGACCTGGGAGGCTTGTGTGACCGACTTACTGGCAGGCGACCTTGGATCCCGAGTGGAGT 5492
QY 1907 TCGTGAACACCCCGCTGCTGTAAGCTGTGTAACAGCTGGGAGAGGAGGCCATCATCG 1966
DB 5493 TCGTGAACACCCCGCTGCTGTAAGCTGTGTAACAGCTGGGAGAGGAGGCCATCATCG 5552
QY 1967 GCGCCGAGACCTTCTACGTGACCGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCG 2026
DB 5553 GCGCCGAGACCTTCTACGTGACCGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCG 5612
QY 2027 GCTACGTGACCGCGCGCGCGCGCGAGAGATCTGTGACCTGACCGAGACCAACACAGA 2086
DB 5613 GCTACGTGACCGCGCGCGCGCGCGAGAGTGTGTGCTGACCGAGACCAACACAGA 5672
QY 2087 AGACCGAGCTGACCGCGCATCCAGCTGCGCTGACAGACAGCGCGAGCTGAACATCG 2146
DB 5673 AGACCGAGCTGACCGCGCATCCAGCTGCGCTGACAGACAGCGCGCTGAGGTGAACATCG 5732
QY 2147 TGACCGAGACCGACCTGACCGCTGGGCATCATCCAGCGCGCGCGAGCAAGAGCGAGAGCG 2206
DB 5733 TGACCGAGACCGACCTGACCGCTGGGCATCATCCAGCGCGCGCGAGCAAGAGCGAGAGCG 5792
QY 2207 AGCTGTGACCGAGATCATCGAGCAGCTGATCAAGAGGAGAGGCTGTACCTGAGCTGGG 2266
DB 5793 AGCTGTGACCGAGATCATCGAGCAGCTGATCAAGAGGAGAGGCTGTACCTGAGCTGGG 5852
QY 2267 TGCCCGCCCAAGGCGATCGCGCGCAACAGCAGATCTCAAGCTGTGTGACCAAGGCGCA 2326
DB 5853 TGCCCGCCCAAGGCGATCGCGCGCAACAGCAGATCTCAAGCTGTGTGACCGCGCGCA 5912
QY 2327 TCCGCAAGGTGCTTCTCTGACCGCATCGATGCGCGCATCGTGTATCTACAGTA 2381
DB 5913 TCCGCAAGGTGCTTCTCTGACCGCATCGATGCGCGCATCGCAAGGCGCGCGAGGAGCGAGAA 5967

RESULT 14
AX427934
LOCUS
DEFINITION 172 from Patent WO0232943.
ACCESSION AX427934
VERSION AX427934.1 GI:21538021
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
AUTHORS artificial sequences.
REFERENCE 1
Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
JOURNAL genetic immunization
GOVERNMENT OF THE UNITED STATES (US) Patent: WO 0232943-A 172 25-APR-2002;
FEATURES Location/Qualifiers
source 1..9783
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 2339 a 2897 c 2723 g 1824 t
ORIGIN /note="plasmid pVR1012x/s containing HIV genes"

Query Match 80.0%; Score 1966.6; DB 6; Length 9783;
Best Local Similarity 90.7%; Pred. No. 3.9e-197;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

QY 14 TGGCGAGGCGCATGAGCAGG--CCACCAGGCGCAACATCTGTGATGAGCGCAGCAACT 70
DB 2964 TGGCGAGGCGCATGAGCAGGTTGACCAAGCGCCACCATTATGATGAGCGCGCAACT 3023
QY 71 TCAAGGGCCCCAAGGCGCATCAAGTCTTCAACTGCGGCAAGAGGCGCCACATCGGCC 130
DB 3024 TCCGCAACAGCGCAAGATCGTGAAGTCTTCAACTGCGGCAAGAGGCGCCACACCGGCC 3083
QY 131 GCAACTGCGGCGCCCCCGCAAGAGGCGTGTGAAATGCGGCAAGAGGCGCCACCAGA 190
DB 3084 GCAACTGCGGCGCCCCCGCAAGAGGCGTGTGAAATGCGGCAAGAGGCGCCACCAGA 3143
QY 191 TGAAGGACTGCACCGAGCGCCAGGCAACTTCTTCGCGAGGACCTGGCCTTCCCGCAGG 250
DB 3144 TGAAGGACTGCACCGAGCGCAGGCTAA-----TAGGAGAGATCTGGCCTTCCACAAAG 3197
QY 251 GCAAGGCCGCGAGTTCCCGCAGCAGAGAGAACCGCGCCCAAGAGCGCCACCAGCGCAGC 310
DB 3198 GGAAGGCCAGGGAATTTTCTTCAGAGCAGACAGAGCCCAACAGCCCCACAGAGAGAGC 3257
QY 311 TGCAGTTCGCGG-----CGACAAACCCCGCAGGAGCGCGGCGCCAGCGCCAGGCA 364
DB 3258 TTCAGGTTTGGGGAGAGAGAACCAACTCCCTCTCAGAAAGCAGGAGCGCCATAGACAAGGAA 3317
QY 365 -----CCTCGAATCTCCCGCAGATCACCTGTGCGACGCGCCCTGTGTGAGCATCAGG 418
DB 3318 CTGTATCTTTAGCTTCCTCAGATCACTTTTGGCAGCAGCCCTCGTCACAATAAGA 3377
QY 419 TGGCGGCGCAGATCAAGAGGCGCCTGTGTGACACCGCGCGCAGACACCTGTGTGGAGG 478
DB 3378 TAGGGGCGCAGCTGAAGGAGGCCCTTCTAGACACCGGCGCGCAGACACCTGTGTGGAG 3437
QY 479 AGATGAGCCTGCCCGGCAAGTTGGAAGCCCAAGATGATGCGGCGGATCGGGCGCTTCATCA 538
DB 3438 AGATGAACTTCCCGGCGCTGGAAGCCCAAGATGATGCGGCGGATCGGGCGCTTCATCA 3497
QY 539 AGGTGCGCAGTACGACAGATCTCTGTGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
DB 3498 AGGTGGGCCAGTACGACAGATCTGTGATGAGATCTGCGGCGCAGAGCCATCGGCACCG 3557
QY 599 TGCTGATCGGCGCCACCGCCGTGAACATCATPCGCGCGCAACATGCTGACCCAGCTGGGCT 658
DB 3558 TGCTGGTGGCGCCACCGCCGTGAACATCATPCGCGCGCAACATGCTGACCCAGATCGGCT 3617
QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGCCTGTAAGCTGAAGCGCGGCA 718
DB 3618 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGCCTGTAAGCTGAAGCGCGGCA 3677
QY 719 TGAAGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGGAGGAAGATCAAGSCCTTGACC 778
DB 3678 TGAAGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGGAGGAAGATCAAGSCCTTGTTGG 3737
QY 779 CCATCTCGGAGGAGATGGAGAGGCGCAAGATCAACCAAGATCGGCCCGCGAGAACCCCT 838
DB 3738 AGATCTGCACCGAGATGGAGAGGAGGCAAGATCAGCAAGATCGGCCCGCGAGAACCCCT 3797
QY 839 ACAACACCCCGTTCGCCATCAAGAGAGGACACACCAAGTGGCGCAAGCTGGTGG 898
DB 3798 ACAACACCCCGTTCGCCATCAAGAGAGGAGACACCAAGTGGCGCAAGCTGGTGG 3857
QY 899 ACTTCCGAGAGTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGTGGGCATCCCC 958
DB 3858 ACTTCCGAGAGTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGTGGGCATCCCC 3917
QY 959 ACCCGCGCGGCTGAAGAAGAAGAGCGTGAACCGTGTGAGCGTGGGCGACGCCCTACT 1018

DB 3918 ACCCGCGCGGCTGAAGCAGAAGAGAGCGTACCGTGTGAGCGTGGCGCAGCGCTACT 3977
QY 1019 TCAGGCTGCCCTGGACGAGGACTTCCCGCAAGTACACCGCCTTCACCATCCCGCAGCATCA 1078
DB 3978 TCAGGCTGCCCTGGACAGGACTTCCCGCAAGTACACCGCCTTCACCATCCCGCAGCATCA 4037
QY 1079 ACAACGAGACCCCGGCATCCGCTACAGTACAAAGTGTGCCCCAGGGCTGGAAGGGA 1138
DB 4038 ACAACGAGACCCCGGCATCCGCTACAGTACAAAGTGTGCCCCAGGGCTGGAAGGGA 4097
QY 1139 GCCCAGCATCTTCAGAGCAGCATGACCAAGATCTGAGAGCCCTTCGCCAAGAGAAC 1198
DB 4098 GCCCAGCATCTTCAGAGCAGCATGACCAAGATCTGAGAGCCCTTCGCCAAGAGAAC 4157
QY 1199 CCGAGATCGTGTATCTACCA-----GGCCCCCTGTGTGAGGAGCAGCTGGAGATCG 1252
DB 4158 CCGAGATCGTGTATCTACCAAGTACATGGACCACTGTGTGAGGAGCAGCTGGAGATCG 4217
QY 1253 GCCAGCACCGGCCCAAGATCGAGGAGCTGCGCAAGACCTGTGTGCGCTGGGCTTACCA 1312
DB 4218 GCCAGCACCGGCCCAAGATCGAGGAGCTGCGCAAGACCTGTGTGCGCTGGGCTTACCA 4277
QY 1313 CCCCAGCAAGAAGCACCAGAGGAGCGCCCTTCTGCCCCAT-----CGAGCTGCAC 1366
DB 4278 CCCCAGCAAGAAGCACCAGAGGAGCGCCCTTCTGTTGGATGGGCTACGAGCTGCAC 4337
QY 1367 CCGCAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAAGAGAGAGCTGGACCGTGAACG 1426
DB 4338 CCGCAAGTGGACCGTGCAGCCCATCGTGTGCCCGAGAAGACAGCTGGACCGTGAACG 4397
QY 1427 ACATCCAGAGCTGTGGGCAAGCTGAACTTGGGCGCAGCAGATCTACCCCGGCATCAAG 1486
DB 4398 ACATCCAGAGCTGTGGGCAAGCTGAACTTGGGCGCAGCAGATCTACCGCGCATCAAG 4457
QY 1487 TGCGCAGCTGTGAAGTGTGTGCGGCGCCAGAGCGCTGACCCACATCTGTCGCCCTCA 1546
DB 4458 TGCGCAGCTGTGAAGTGTGTGCGGCGCCAGAGCGCTGACCCAGCTGTGTCGCCCTCA 4517
QY 1547 CCGAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGCGCGAGCGCCGTCAGC 1606
DB 4518 CCGAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGAGAGGAGCGCCGTCAGC 4577
QY 1607 CGGTGTACTACGACCCAGCAGGAGCTGTGTGGCGCAGATCTCAGAGAGGCGGCCAGAC 1666
DB 4578 CGGTGTACTACGACCCAGCAGGAGCTGTGTGGCGCAGATCTCAGAGAGGCGGCCAGAC 4637
QY 1657 AGTGACCTTACAGATCTTACAGGAGCGCTTCAAGAACTGAAAGACCGGCAAGTACGCCA 1726
DB 4638 AGTGACCTTACAGATCTTACAGGAGCGCTTCAAGAACTGAAAGACCGGCAAGTACGCC 4697
QY 1727 AGATGCGCACCGCCACACCAAGCAGTGAAGCAGCTGACCGAGCGCGTGCAGAGATCG 1786
DB 4698 GCATGAAGGCGCGCCACACCAAGCAGTGAAGCAGCTGACCGAGCGCGTGCAGAGATCG 4757
QY 1787 CCATGAGAGAGATCGTGTATCTGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAGG 1846
DB 4758 CCACCGAGAGATCGTGTATCTGGGCAAGACCCCAAGTTCAAGCTGCCATCCAGAGG 4817
QY 1847 AGACTGGAGACCTTGTGTGAGCCGACTACTGCGAGCGCCACCTGGATCCCGAGTGGGAGT 1906
DB 4818 AGACTGGAGACCTTGTGTGAGCCGACTACTGCGAGCGCCACCTGGATCCCGAGTGGGAGT 4877
QY 1907 TCGTGAACACCCCGCTTGTGTGAGCTGTGTGTACAGCTGGAGAGAGCGCCATCATCG 1966
DB 4878 TCGTGAACACCCCGCTTGTGTGAGCTGTGTGTACAGCTGGAGAGAGCGCCATCATCG 4937
QY 1967 GCGCGAGAGCTTCTAGCTGAGCGCGCCCAAGCGGAGACCAAGATCGCGCAGGCGCG 2026
DB 4938 GCGCGAGAGCTTCTAGCTGAGCGCGCGCCCAAGCGGAGACCAAGTGGGCAAGGCGCG 4997
QY 2027 GCTAGCTGACCGACCGGCGCGCAGAGATCTGTGAGCTGACCGAGACCAACCAAGAG 2086

Db 4998 GCTACGTGACCGACCGCGCGCCAGAGGTGGTGGCCCTGACCGACACCAACACAGA 5057
Qy 2087 AGACCGAGCTGACGAGCATCCAGCTGGCCCTGCAGGACGCGCAGGAGTGAACATCG 2146
Db 5058 AGACCGAGCTGACGAGCATCCAGCTGGCCCTGCAGGACGCGCAGGAGTGAACATCG 5117
Qy 2147 TGACCGACAGCCAGTACGCGCTGGCATCATCCAGGCGCCAGCCGACAAAGAGGAGCG 2206
Db 5118 TGACCGACAGCCAGTACGCGCTGGCATCATCCAGGCGCCAGCCGACAAAGAGGAGCG 5177
Qy 2207 AGCTGGTGAACAGATCATCGACAGCTGATCAAGAGGAGAGGTACTGAGCTGGG 2266
Db 5178 AGCTGGTGAACAGATCATCGACAGCTGATCAAGAGGAGAGGTACTGAGCTGGG 5237
Qy 2267 TGCCCGCCCAACAGGAGCGCGGCGCAACGACGATCGACAGCTGCTGAGCAAGGCGCA 2326
Db 5238 TGCCCGCCCAACAGGAGCGCGGCGCAACGACGAGTGGAGCGGCTGGTGGCGCGGCA 5297
Qy 2327 TCCGCAAGGTGCTTCTTCCAGCGGCATCGATGGCGGCATCGTGAATACCACTA 2381
Db 5298 TCCGCAAGGTGCTTCTTCCAGCGGCATCGACAAAGGCCAGGAGGACACAGAA 5352

RESULT 15
AX427932 AX427932 9792 bp DNA linear PAT 20-JUN-2002
LOCUS Sequence 170 from Patent W00232943.
DEFINITION AX427932
ACCESSION AX427932
VERSION AX427932.1 GI:21538019
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 170 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
source 1..9792
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012/s containing HIV genes"

BASE COUNT 2339 a 2898 c 2725 g 1830 t
ORIGIN

Query Match 80.08; Score 1966.6; DB 6; Length 9792;
Best Local Similarity 90.78; Pred. No. 3.9e-197;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

Qy 14 TGGCGGAGGCCATGAGCCAGG---CCACGAGCGCAACATCTGATGCGGCGAGCAACT 70
Db 2967 TGGCGGAGGCCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGCGGCGGCAACT 3026
Qy 71 TCAGAGGCCCCAAGCGCATCATCAAGTCTTCACTGCGGCAAGGAGGCGGCACATCGGCC 130
Db 3027 TCCGCAACAGCGCAAGATCGTGAAGTCTTCACTGCGGCAAGGAGGCGGCACACCGGCC 3086
Qy 131 GCAACTGCCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACCA 190
Db 3087 GCAACTGCCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACCA 3146
Qy 191 TGAAGGACTGCACCGAGGCGCCAGCCAACTCTTCCGCGGAGGACTGGCCCTTCCCGCAGG 250
Db 3147 TGAAGGACTGCACCGAGGCGCCAGCCAACTCTTCCGCGGAGGACTGGCCCTTCCCGCAGG 3200
Qy 251 GCAAGGCGCGAGTTCCTCCAGCGAGAGAAACCGCCCAACAGCCCCACAGCGCGGAGC 310
Db 3201 GGAAGGCGCGAGTTCCTCCAGCGAGAGAAACCGCCCAACAGCCCCACAGCGCGGAGC 3260
Qy 311 TGCAGGTGCGGG-----CGACAAACCCCGAGGAGGCGGCGCGGAGGCGCGGCA 364

Db 3261 TTCAAGTTTGGGGAAGAGACAACAACATCCCTCTCAGAAAGAGGAGCGCATAGACAAGAA 3320
Qy 365 -----CCCTGAACCTTCCCGCAGATCACCTTGTGGCAGCGCCCTGGTGGAGCATCAAG 418
Db 3321 CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCGACCCCTGTCACATAAGA 3380
Qy 419 TGGCGCGCCAGATCAAGGAGCCCTGCTGACACACCGCGCGGCGAGCAGCACCTGCTGGAG 478
Db 3381 TAGGGGCGCAGCTGAAGAGGCGCCTTCTAGACACCGCGCGCGAGACACCTGCTGGAGG 3440
Qy 479 AGATGACCTTCCCGCGCAAGTGGAAAGCCCAAGATGATCGGGGCGATCGGGGCTTCATCA 538
Db 3441 AGATGACCTTCCCGCGCGCTGGAAAGCCCAAGATGATCGGGGCGATCGGGGCTTCATCA 3500
Qy 539 AGGTGCCCGCAGTACACACAGATCTGATCGAGATCTCGGCAAGAAGGCCATCGGCACCG 598
Db 3501 AGGTGCCCGCAGTACACACAGATCTGATCGAGATCTCGGCAAGAAGGCCATCGGCACCG 3560
Qy 599 TGTGATCGGCGCCACCCCGGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT 658
Db 3561 TGTGATCGGCGCCACCCCGGTGAACATCATCGGCGCGCAACATGCTGACCCAGATCGGCT 3620
Qy 659 GCACCTGAACTTCCCATCAGCGCCCATCGAGACCGTGCCTGGAAGTGAAGCCCGGCA 718
Db 3621 GCACCTGAACTTCCCATCAGCGCCCATCGAGACCGTGCCTGGAAGTGAAGCCCGGCA 3680
Qy 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGGAGGAGAAGATCAAGCCCTGACCG 778
Db 3681 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGGAGGAGAAGATCAAGCCCTGAGG 3740
Qy 779 CCATCTCGGAGGAGTGGAGAGGAGGCAAGATCACCAAGATCGGCGCGCGAGAACCCCT 838
Db 3741 AGATCTGCACCGAGATGGAGAGGAGGCAAGATCAGCAAGATCGGCGCGCGAGAACCCCT 3800
Qy 839 ACACACCCCGCTTCCCATCAAGAGAGGAGCAGCACCACCAAGTGGCGCAAGCTGGTGG 898
Db 3801 ACACACCCCGCTTCCCATCAAGAGAGGAGCAGCACCACCAAGTGGCGCAAGCTGGTGG 3860
Qy 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTGCAGTGGGCGATCCGCC 958
Db 3861 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTGCAGTGGGCGATCCGCC 3920
Qy 959 ACCCGCGCGCTGAAGAAGAGAGAGCGTGCAGGTGCTGGAGCGTGGCGCGACGCTACT 1018
Db 3921 ACCCGCGCGCTGAAGAAGAGAGAGCGTGCAGGTGCTGGAGCGTGGCGCGACGCTACT 3980
Qy 1019 TCAGCGTCCCTTGACGAGGACTTCCCAAGTACACCGCTTACCATCCCCAGCATCA 1078
Db 3981 TCAGCGTCCCTTGACGAGGACTTCCCAAGTACACCGCTTACCATCCCCAGCATCA 4040
Qy 1079 ACAACGAGACCCCGGCATCCGCTACCAAGTACAAGTGTCTGCCCGAGGGCTGGAGGGCA 1138
Db 4041 ACAACGAGACCCCGGCATCCGCTACCAAGTACAAGTGTCTGCCCGAGGGCTGGAGGGCA 4100
Qy 1139 GCGCGAGCATCTTCAGAGAGCATGACCAAGATCTGGAGCCCTTCCCGCGCCGCAACC 1198
Db 4101 GCGCGAGCATCTTCAGAGTGCAGCATGACCAAGATCTGGAGCCCTTCCCGCGCGCAACC 4160
Qy 1199 CCGAGATCGTGTATCTACCA-----GGCGCCCTTGTAGTGGGCGAGCAGCTGGAGATCG 1252
Db 4161 CCGAGATCGTGTATCTACCA-----GGCGCCCTTGTAGTGGGCGAGCAGCTGGAGATCG 4220
Qy 1253 GCCAGCAGCGCGCAAGATCGAGGAGTGCACAGCACTGCTGGCTGGGGCTTTCACCA 1312
Db 4221 GCCAGCAGCGCGCAAGATCGAGGAGTGCACAGCACTGCTGGCTGGGGCTTTCACCA 4280
Qy 1313 CCGCGAGCAAGAGCAGCAGAGGAGGCGCCCTTCTCCGCCAT-----CGAGCTGCACC 1366
Db 4281 CCGCGAGCAAGAGCAGCAGAGGAGGCGCCCTTCTCTGTGATGGGCTACGAGCTGCACC 4340
Qy 1367 CCGACAAGTGGAGCGTGCAGCCCATCGAGCTGCCCGAGAGGAGAGCTGGACCGTGAACG 1426
Db 4341 CCGACAAGTGGAGCGTGCAGCCCATCGAGCTGCCCGAGAGGAGAGCTGGACCGTGAACG 4400

QY 1427 ACATCCAGAAAGCTGGTGGCAAGCTGAACCTGGCCAGCCAGATCTACCCCGGCATCAAGG 1486
|||||
Db 4401 ACATCCAGAAAGCTGGTGGCAAGCTGAACCTGGCCAGCCAGATCTACCCCGGCATCAAGG 4460
QY 1487 TGGCCAGCTGTGCAAGCTGCTGGCGGCGCCAAAGGCCCTGACCGACATCGTGCCCTGA 1546
|||||
Db 4461 TGGCCAGCTGTGCAAGCTGCTGGCGGCGCCAAAGGCCCTGACCGAGGTGGTGCCCTGA 4520
QY 1547 CCGAGGAGCGGAGCTGGAGCTGGCCAGAACCGCGAGATCTTGGCGGAGCCCGTGCAAG 1606
|||||
Db 4521 CCGAGGAGCGGAGCTGGAGCTGGCCAGAACCGCGAGATCTTGGCGGAGCCCGTGCAAG 4580
QY 1607 GCGTGACTACGACCCAGCAAGGACCTGGTGGCGGAGATCCAGAAAGCAGGCGCACGACC 1666
|||||
Db 4581 GCGTGACTACGACCCAGCAAGGACCTGGTGGCGGAGATCCAGAAAGCAGGCGCACGACC 4640
QY 1667 AGTGGACCTACAGATCTACCAAGAGCCCTTCAAGAACCTGAAGACCCGGAAGTACGCCA 1726
|||||
Db 4641 AGTGGACCTACAGATCTACCAAGAGCCCTTCAAGAACCTGAAGACCCGGAAGTACGCCC 4700
QY 1727 AGATGCCACCGCCACACCAAGACGTGAAGCAGCTGACCGAGGCGCGTGCAAGATCG 1786
|||||
Db 4701 GCATGAAGGCGGCCACACCAAGACGTGAAGCAGCTGACCGAGGCGCGTGCAAGATCG 4760
QY 1787 CCATGGAGAGCATCGTGATCTGGGGCAAGACCCCAAGTTCCGGCTGCCCATCCAGAAGG 1846
|||||
Db 4761 CCACCGAGAGCATCGTGATCTGGGGCAAGACCCCAAGTTCAAGCTGCCCATCCAGAAGG 4820
QY 1847 AGACCTGGGAGACCTGGTGGACGACTACTGGCAGGCGCACCTGGATCCCGAGTGGGAGT 1906
|||||
Db 4821 AGACCTGGGAGACCTGGTGGACGACTACTGGCAGGCGCACCTGGATCCCGAGTGGGAGT 4880
QY 1907 TCGTGAACACCCCGCTGGTGAAGCTGTGGTACAGCTGGAGAGGAGCGCCATCATCG 1966
|||||
Db 4881 TCGTGAACACCCCGCTGGTGAAGCTGTGGTACAGCTGGAGAGGAGCGCCATCATCG 4940
QY 1967 GCGCCGAGACCTTCTAGCTGGAGCGCGCCCAACCGCGAGACCAAGATCGCAAGGCCG 2026
|||||
Db 4941 GCGCCGAGACCTTCTAGCTGGAGCGCGCCCAACCGCGAGACCAAGATCGCAAGGCCG 5000
QY 2027 GCTACGTGACCGACCGCGCGCGCAGAGATCGTGAAGCTGACCGTGAAGACCAACCAAG 2086
|||||
Db 5001 GCTACGTGACCGACCGCGCGCGCAGAGTGGTGGCCCTGACCGACACCAACCAAG 5060
QY 2087 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACACGCGGCGAGGTGAACATCG 2146
|||||
Db 5061 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACACGCGGCGTGGAGGTGAACATCG 5120
QY 2147 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCCGACAAAGAGCGAGAGCG 2206
|||||
Db 5121 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCCGACAAAGAGCGAGAGCG 5180
QY 2207 AGGTGTGAACAGATCATCGACAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 2266
|||||
Db 5181 AGGTGTGAGCCAGATCATCGACAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 5240
QY 2267 TSCCGCCCAACAGGCGATCGGCGGCAACAGCAGATCGACAGCTGGTGAAGCAAGGCCA 2326
|||||
Db 5241 TSCCGCCCAACAGGCGATCGGCGGCAACAGCAGAGTGGACGGCCTGGTGAAGCGCGGCA 5300
QY 2327 TCCGCAAGGTGCTTCTCGGACGGCATCGATGGCGGCATCGTGTATCTACAGTA 2381
|||||
Db 5301 TCCGCAAGGTGCTTCTCGGACGGCATCGACAAAGGCCAGGAGGAGCACAGAA 5355

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 10:51:03 ; Search time 489.804 Seconds

(without alignments)
11296.684 Million cell updates/sec

Title: US-09-610-313-32

Perfect score: 2457

Sequence: 1 gtcgacgccaccatggccga.....gggctagcaccgggaattc 2457

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2457	100.0	2457	24	ABL39961
2	2436.2	99.2	2463	24	ABL39960
3	2415.4	98.3	2469	24	ABL39959
4	2040	83.0	2300	21	AAA70480
5	2019.2	82.2	2306	21	AAA70479
6	2000.8	81.4	9166	24	ABK91616
7	1998.4	81.3	2312	21	AAA70481
8	1992.4	81.1	9788	24	ABK91622
9	1978.4	80.5	9169	24	ABK91617

10	1975	80.4	9194	24	ABK91612	Modified HIV prote
11	1973.4	80.3	9194	24	ABK91611	Modified HIV prote
12	1973.4	80.3	12411	24	ABK91613	Modified HIV prote
13	1970	80.2	9785	24	ABK91624	Modified HIV prote
14	1966.6	80.0	9167	24	ABK91619	Modified HIV prote
15	1966.6	80.0	9170	24	ABK91614	Modified HIV prote
16	1966.6	80.0	9782	24	ABK91621	Modified HIV prote
17	1966.6	80.0	9783	24	ABK91620	Modified HIV prote
18	1966.6	80.0	9792	24	ABK91618	Modified HIV prote
19	1965	80.0	9189	24	ABK91607	Modified HIV prote
20	1945.4	79.2	3015	24	ABL39983	Synthetic Pol poly
21	1928.8	78.5	3009	24	ABL40024	HIV Gag-polymerase
22	1915.8	78.0	4319	21	AAA70414	HIV Gag-polymerase
23	1882.2	76.6	8908	21	RAZ52055	Synthetic protease
24	1878.6	76.5	1965	24	ABL39989	Packaging construc
25	1872.8	76.2	2299	21	AAA70478	Synthetic protease
26	1859.8	75.7	2577	24	RAA141592	HIV FS(+)-Proteinac
27	1852.2	75.4	4343	24	RAI67896	HIV-1 subtype C is
28	1852	75.4	2305	21	AAA70477	Nucleotide sequenc
29	1848.2	75.2	4341	24	RAI67897	HIV FS(+)-Proteinac
30	1827.6	74.4	1978	24	ABL39991	Nucleotide sequenc
31	1825.6	74.3	7897	24	ABK91609	Synthetic protease
32	1762	71.7	3012	21	RAZ52054	Modified HIV prote
33	1757.2	71.5	9407	24	ABK91623	Codon optimised Hu
34	1607.4	65.4	4307	20	AAZ08740	Modified HIV prote
35	1607.4	65.4	4307	21	AAA93972	HIV gagpol-SVNgp C
36	1607.4	65.4	4307	22	AAC86876	Human immunodefici
37	1599.4	65.1	4307	22	AAH43696	Nucleotide sequenc
38	1599.4	65.1	4327	21	AAA93984	Codon optimised HI
39	1599.4	65.1	4353	21	AAA93983	HIV partial leader
40	1599.4	65.1	4642	21	AAA93982	HIV partial leader
41	1599.4	65.1	9772	22	AAH43699	HIV complete leade
42	1596	65.0	1668	24	ABL39999	Codon optimised ga
43	1554.4	63.3	1680	24	ABL39997	Synthetic RT polyn
44	1539.2	62.6	4338	22	AAH22806	Synthetic RT polyn
45	1539.2	62.6	8366	22	AAH22810	DNA sequence of a

ALIGNMENTS

RESULT 1
ABL39961
ID ABL39961 standard; DNA; 2457 BP.
XX
AC ABL39961;
XX
DT 15-MAY-2002 (first entry)
XX
DE Synthetic construct PR975YMWV SEQ ID NO:32.
XX
KW Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX
OS Human immunodeficiency virus type C.
OS Synthetic.
XX
PN WO200204493-A2.
XX
PD 17-JAN-2002.
XX
PF 05-JUL-2001; 2001WO-US21241.
XX
PR 05-JUL-2000; 2000US-0610313.
XX
PA (CHIR) CHIRON CORP.
PA Synthetic construc
(UYST-) UNIV STELLENBOSCH.
XX
XX HIV FS(-)_ProtMod.
PI Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX HIV FS(-)_ProtMod.
DR WPI; 2002-154920/20.
XX

PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
PS cell lines, particularly in gene therapy -

XX Claim 1; Fig 10; 233pp; English.

CC The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
CC Nef (1). (1) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV Type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the
CC exemplification of the present invention.

XX Sequence 2457 BP; 566 A; 837 C; 754 G; 300 T; 0 other;

Query Match 100.0%; Score 2457; DB 24; Length 2457;

Best Local Similarity 100.0%; Pred. No. 1.9e-297;

Matches 2457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACGCCACCATGCGCGAGGCCATGAGCCAGGCCACCAAGCTTCAACTGCGCGAAGGGGC 120
DB 1 GTGACGCCACCATGCGCGAGGCCATGAGCCAGGCCACCAAGCTTCAACTGCGCGAAGGGGC 60
QY 61 CGCAGCAACTTCAAGGGCCCCAGCGCATCATCAAGTGTTCACCTGCGCGAAGGGGC 120
DB 61 CGCAGCAACTTCAAGGGCCCCAGCGCATCATCAAGTGTTCACCTGCGCGAAGGGGC 120
QY 121 CATCATCGCCCGCAACTGCCCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAG 180
DB 121 CATCATCGCCCGCAACTGCCCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAG 180
QY 181 GGGCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCGCGAGGACCTGGCC 240
DB 181 GGGCACCAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCGCGAGGACCTGGCC 240
QY 241 TTCCCCCGAGGCGAGGCGCGAGTTCCTCCAGCGAGCAGCAACCGCGCAACACGCCCCACC 300
DB 241 TTCCCCCGAGGCGAGGCGCGAGTTCCTCCAGCGAGCAGCAACCGCGCAACACGCCCCACC 300
QY 301 AGCGCGAGCTGCAAGTGCAGCGCGCAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCG 360
DB 301 AGCGCGAGCTGCAAGTGCAGCGCGCAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCG 360
QY 361 GGCACCTGAACCTTCCCGCAGATCACTGTCGAGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB 361 GGCACCTGAACCTTCCCGCAGATCACTGTCGAGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 421 GCGCGCCAGATCAAGGAGGCGCTTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 GCGCGCCAGATCAAGGAGGCGCTTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 ATGAGCTGCGCGCGAGTGGAGCGCCCAAGATGATCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 ATGAGCTGCGCGCGAGTGGAGCGCCCAAGATGATCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 GTGCGCGCAGTACGACCGATTCCTGTAGATCTGCGCGCAAGAGGCCATCGCGCACCGTG 600
DB 541 GTGCGCGCAGTACGACCGATTCCTGTAGATCTGCGCGCAAGAGGCCATCGCGCACCGTG 600
QY 601 CTGATCGGCG 660
DB 601 CTGATCGGCG 660
QY 661 ACCCTGAACCTTCCCGCAGTACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 ACCCTGAACCTTCCCGCAGTACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 GACGGCCCCAAGGTGAAGCAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780

DB 721 GACGGCCCCAAGGTGAAGCAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
QY 781 ATCTGCGAGGAGATGGAGAGGAGGCAAGATCACCAGATCGCGCGCGCGCGCGCGCGCGCG 840
DB 781 ATCTGCGAGGAGATGGAGAGGAGGCAAGATCACCAGATCGCGCGCGCGCGCGCGCGCGCG 840
QY 841 AACACCCCCCTGTTCGCCATCAAGAAGAGAGCAGCACCAGTGGCGCAAGCTGGTGGAC 900
DB 841 AACACCCCCCTGTTCGCCATCAAGAAGAGAGCAGCACCAGTGGCGCAAGCTGGTGGAC 900
QY 901 TTCGGGAGCTGAACAGCCACCCAGGACTTCTGGAGGTGAGCTGGGAGTATCCCCAC 960
DB 901 TTCGGGAGCTGAACAGCCACCCAGGACTTCTGGAGGTGAGCTGGGAGTATCCCCAC 960
QY 961 CCCGCGCGCTGAAGAGAGAGAGCGTGACCGTGTGACGTGGCGCGCGCGCGCGCGCGCG 1020
DB 961 CCCGCGCGCTGAAGAGAGAGAGCGTGACCGTGTGACGTGGCGCGCGCGCGCGCGCGCG 1020
QY 1021 AGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCTCCCGCAGCATCAAC 1080
DB 1021 AGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCTCCCGCAGCATCAAC 1080
QY 1081 AACGAGACCCCCCGCATCCGCTACAGTACAAAGTGTGCTGCCCGCAGGCGTGAAGGGCAG 1140
DB 1081 AACGAGACCCCCCGCATCCGCTACAGTACAAAGTGTGCTGCCCGCAGGCGTGAAGGGCAG 1140
QY 1141 CCAGAGCTTTCAGAGCAGCAGTACCAAGATCTGAGGCGCTTCCGCGCGCGCGCGCGCGCG 1200
DB 1141 CCAGAGCTTTCAGAGCAGCAGTACCAAGATCTGAGGCGCTTCCGCGCGCGCGCGCGCGCG 1200
QY 1201 GAGTGTGTATCTACAGGCGCGCGCTGTAGTGTGGCGAGCAGCTTGGAGATCGGCGCAG 1260
DB 1201 GAGTGTGTATCTACAGGCGCGCGCTGTAGTGTGGCGAGCAGCTTGGAGATCGGCGCAG 1260
QY 1261 CGCGCCAGATCGAGGAGCTGGCGAAGCAGCTGCTGCGCTGGGCGCTTCCACCGCGCGCG 1320
DB 1261 CGCGCCAGATCGAGGAGCTGGCGAAGCAGCTGCTGCGCTGGGCGCTTCCACCGCGCGCG 1320
QY 1321 AAGAAGCACCAGAGGAGCGCGCGCTTCTGCGCATCGAGCTGACCGCGCGCGCGCGCGCGCG 1380
DB 1321 AAGAAGCACCAGAGGAGCGCGCGCTTCTGCGCATCGAGCTGACCGCGCGCGCGCGCGCGCG 1380
QY 1381 GTGAGCGCTGAGCTGCG 1440
DB 1381 GTGAGCGCTGAGCTGCG 1440
QY 1441 GTGGGCAAGCTGAAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
DB 1441 GTGGGCAAGCTGAAGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
QY 1501 AAGTGTGTGCG 1560
DB 1501 AAGTGTGTGCG 1560
QY 1561 CTGAGCTGCG 1620
DB 1561 CTGAGCTGCG 1620
QY 1621 CCAGCAGAGACCTGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1680
DB 1621 CCAGCAGAGACCTGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1680
QY 1681 ATCTTACCGAGGCGCTTCAAGAAGCTTCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1740
DB 1681 ATCTTACCGAGGCGCTTCAAGAAGCTTCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1740
QY 1741 CACACCAACAGCTGAAGCAGCTGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1800
DB 1741 CACACCAACAGCTGAAGCAGCTGACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1800
QY 1801 GTGATCTGGGCG 1860

Db 1801 GTGATCTGGGCGAAGACCCCAAGTTCGCGCTGCCATCCAGAAGGAGACCTGGGAGACC 1860
Qy 1861 TGGTGGACCGACTACTGCGAGCCACCTGGATCCCGAGTGGAGTTCGTGAACACCCCC 1920
Db 1861 TGGTGGACCGACTACTGCGAGCCACCTGGATCCCGAGTGGAGTTCGTGAACACCCCC 1920
Qy 1921 CCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGCGCGAGACCTTC 1980
Db 1921 CCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGCGCGAGACCTTC 1980
Qy 1981 TACGTGGACGGCGCCCAACCGCAGACCAAGATCGGAAGCGCGGTACGTGACCGAC 2040
Db 1981 TACGTGGACGGCGCCCAACCGCAGACCAAGATCGGAAGCGCGGTACGTGACCGAC 2040
Qy 2041 CGGGCCGCGCAGAAAGATCGTAGCCTGACCGAGACCAACCAAGAACCGAGCTGCAG 2100
Db 2041 CGGGCCGCGCAGAAAGATCGTAGCCTGACCGAGACCAACCAAGAACCGAGCTGCAG 2100
Qy 2101 GGCATTCAGCTGGCCCTGCAGACAGCGGACGAGGTGAACATCGTGACCGACACCGAG 2160
Db 2101 GGCATTCAGCTGGCCCTGCAGACAGCGGACGAGGTGAACATCGTGACCGACACCGAG 2160
Qy 2161 TACGCCCTGGGCATCATCCAGGCCAGCCCGACAGAGCGAGCGAGCTGGTGAACACG 2220
Db 2161 TACGCCCTGGGCATCATCCAGGCCAGCCCGACAGAGCGAGCGAGCTGGTGAACACG 2220
Qy 2221 ATCATCGAGCAGCTGATCAAGAAGAGAGAGTGTACCTGAGCTGGTGGCGCCCAACAG 2280
Db 2221 ATCATCGAGCAGCTGATCAAGAAGAGAGAGTGTACCTGAGCTGGTGGCGCCCAACAG 2280
Qy 2281 GGCATCGCGCGCAGCAGATCGACAAAGCTGGTGAAGAGGGGCATCCGCAAGGTGCTG 2340
Db 2281 GGCATCGCGCGCAGCAGATCGACAAAGCTGGTGAAGAGGGGCATCCGCAAGGTGCTG 2340
Qy 2341 TTCTTGGACGCGATCGATGGCGGCATCGTATCTACCACTACATGAGCAGACCTGTACGTG 2400
Db 2341 TTCTTGGACGCGATCGATGGCGGCATCGTATCTACCACTACATGAGCAGACCTGTACGTG 2400
Qy 2401 GGCAGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2457
Db 2401 GGCAGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2457

RESULT 2
ABL39960
ID ABL39960 standard; DNA; 2463 BP.
AC ABL39960;
XX
AC
XX
DT 15-MAY-2002 (first entry)
XX
DE Synthetic construct PR975YM SEQ ID NO:31.
XX
KW Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX
OS Human immunodeficiency virus type C.
OS Synthetic.
OS
PN W0200204493-A2.
XX
PD 17-JAN-2002.
XX
XX 05-JUL-2001; 2001WO-US21241.
XX
XX 05-JUL-2000; 2000US-0610313.
XX
XX (CHIR) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX
XX Zur Wegede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
PI
XX

DR WPI; 2002-154920/20.
XX
PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
XX cell lines, particularly in gene therapy
XX
PS Claim 1; Fig 9; 233pp; English.
XX
CC The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
CC Nef (1). (1) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the
XX exemplification of the present invention.
SQ Sequence 2463 BP; 567 A; 835 C; 759 G; 302 T; 0 other;

Query Match 99.2%; Score 2436.2; DB 24; Length 2463;
Best Local Similarity 99.6%; Pred. No. 7.2e-295;
Matches 2454; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

Qy 1 GTCGAGCGCCACCATGCGCGAGGCCCATGAGCCAGGCCACCGCCCAACATCTGTATGCAG 60
Db 1 GTCGAGCGCCACCATGCGCGAGGCCCATGAGCCAGGCCACCGCCCAACATCTGTATGCAG 60
Qy 61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTCTTCAACTTCGCGCAAGGAGGCG 120
Db 61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTCTTCAACTTCGCGCAAGGAGGCG 120
Qy 121 CACATCGCGCGCACTGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTTCGCGCAAGGAG 180
Db 121 CACATCGCGCGCACTGCGCGCCCGCCCGCAAGAGGGCTGCTGGAAGTTCGCGCAAGGAG 180
Qy 181 GGCACACAGATGAAGAGTGCACCGCGCCGCGCCCAACTTCTTCGCGAGGACCTGGCC 240
Db 181 GGCACACAGATGAAGAGTGCACCGCGCCGCGCCCAACTTCTTCGCGAGGACCTGGCC 240
Qy 241 TTCCCCCAGGCGAAGGCCCGCGAGTTCCCGACGAGAGAGCCGGCCCAACAGCCCGCCAC 300
Db 241 TTCCCCCAGGCGAAGGCCCGCGAGTTCCCGACGAGAGAGCCGGCCCAACAGCCCGCCAC 300
Qy 301 AGCCCGGAGCTGCAGGTGGCGGGCGACAAACCCCGCGAGGAGCGCGCGCGAGCCCGAG 360
Db 301 AGCCCGGAGCTGCAGGTGGCGGGCGACAAACCCCGCGAGGAGCGCGCGCGAGCCCGAG 360
Qy 361 GGCACCTGAACTTCCCGCAGATCACCTGTGCGAGCGCCCGCTGTGAGCATCAAGGTG 420
Db 361 GGCACCTGAACTTCCCGCAGATCACCTGTGCGAGCGCCCGCTGTGAGCATCAAGGTG 420
Qy 421 GCGGCCAGATCAAGAGGCGCTGCTGGACACCGCGCGCGAGACACCGTGTGAGGAG 480
Db 421 GCGGCCAGATCAAGAGGCGCTGCTGGACACCGCGCGCGAGACACCGTGTGAGGAG 480
Qy 481 ATGAGCTTCCCGCGCAAGTGAAGCCCAAGATGATGCGCGGATCGCGGGCTTCATCAAG 540
Db 481 ATGAGCTTCCCGCGCAAGTGAAGCCCAAGATGATGCGCGGATCGCGGGCTTCATCAAG 540
Qy 541 GTGGCCAGTACGACCATCTGATCGAGATCTCGGCAAGAGCCCATCGGCAACCGTG 600
Db 541 GTGGCCAGTACGACCATCTGATCGAGATCTCGGCAAGAGCCCATCGGCAACCGTG 600
Qy 601 CTGATCGCGCCACCGCGCTGAACATCATCGCGCCCAACATGCTGACCCAGCTGGGCTGC 660
Db 601 CTGATCGCGCCACCGCGCTGAACATCATCGCGCCCAACATGCTGACCCAGCTGGGCTGC 660
Qy 661 ACCCTGAAGTTCGCCATCAGCCCGCATCGAGACCGGTGCCGTGAAGCTGAAGCCCGCATG 720
Db 661 ACCCTGAAGTTCGCCATCAGCCCGCATCGAGACCGGTGCCGTGAAGCTGAAGCCCGCATG 720

QY 721 GACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCC 780
Db 721 GACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCC 780
QY 781 ATCTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTAC 840
Db 781 ATCTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTAC 840
QY 841 AACACCCCGTGTTCGCCATCAAGAGAAGAGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
Db 841 AACACCCCGTGTTCGCCATCAAGAGAAGAGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
QY 901 TTCGGGAGCTGAACAGCCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCGCCAC 960
Db 901 TTCGGGAGCTGAACAGCCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCGCCAC 960
QY 961 CCGCGCGCTGAAGAGAAGAGAGCGTGACCGTGTGGACGTGGCGACGCTACTTC 1020
Db 961 CCGCGCGCTGAAGAGAAGAGAGCGTGACCGTGTGGACGTGGCGACGCTACTTC 1020
QY 1021 AGGTGCCCTGACGAGGACTTCCGCAAGTACACGCCCTTCACCATCCCGACATCAAC 1080
Db 1021 AGGTGCCCTGACGAGGACTTCCGCAAGTACACGCCCTTCACCATCCCGACATCAAC 1080
QY 1081 AACGAGACCCCGGCATCCGCTACAGTACAAGTGTGCTGCCAGGGCTGGAAGGGCAGC 1140
Db 1081 AACGAGACCCCGGCATCCGCTACAGTACAAGTGTGCTGCCAGGGCTGGAAGGGCAGC 1140
QY 1141 CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGAGACCCCTTCGCGGCCCGCAACCCC 1200
Db 1141 CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGAGACCCCTTCGCGGCCCGCAACCCC 1200
QY 1201 GAGTGTGATCTACAGGCCCGCCCTGTAGCTGGGAGAGCACTTGAGATGGCGCAGC 1260
Db 1201 GAGTGTGATCTACAGGCCCGCCCTGTAGCTGGGAGAGCACTTGAGATGGCGCAGC 1260
QY 1261 CGCGCAAGATCGAGAGCTGCGCAGCAGCCTGTGCTGGCTGGGGCTTCACACCCCGCAG 1320
Db 1261 CGCGCAAGATCGAGAGCTGCGCAGCAGCCTGTGCTGGCTGGGGCTTCACACCCCGCAG 1320
QY 1321 AAGAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1374
Db 1321 AAGAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
QY 1375 TGGACGTCAGCCCATCGAGCTGCCGAGAGAGAGAGTGGACCTGACGACATCCAG 1434
Db 1381 TGGACGTCAGCCCATCGAGCTGCCGAGAGAGAGAGTGGACCTGACGACATCCAG 1440
QY 1435 AAGCTGTGGCAAGCTGAAGTGGGCGAGCAGATCTACCCCGCATCAAGGTGCGCGAG 1494
Db 1441 AAGCTGTGGCAAGCTGAAGTGGGCGAGCAGATCTACCCCGCATCAAGGTGCGCGAG 1500
QY 1495 CTGTGAAGTGTGTGCGGCGGCGCAAGGCCCTGACCGACATCTGCCCCCTGACCGAGGAG 1554
Db 1501 CTGTGAAGTGTGTGCGGCGGCGCAAGGCCCTGACCGACATCTGCCCCCTGACCGAGGAG 1560
QY 1555 GCGAGCTGAGCTGGCGGAGACCGGAGATCTGCGCGAGCCCTGACGCGGCTGTAC 1614
Db 1561 GCGAGCTGAGCTGGCGGAGACCGGAGATCTGCGCGAGCCCTGACGCGGCTGTAC 1620
QY 1615 TAGGACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1674
Db 1621 TAGGACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
QY 1675 TACAGATCTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1734
Db 1681 TACAGATCTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
QY 1735 ACCGCCACACCAAGCAGCTGAGAGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1794
Db 1741 ACCGCCACACCAAGCAGCTGAGAGCTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800

QY 1795 AGCATCGTGTATCTGGGCAAGACCCCAAGTTCGGCTGCCATCCAGAAGGAGACCTGG 1854
Db 1801 AGCATCGTGTATCTGGGCAAGACCCCAAGTTCGGCTGCCATCCAGAAGGAGACCTGG 1860
QY 1855 GAGACCTGTGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAAC 1914
Db 1861 GAGACCTGTGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAAC 1920
QY 1915 ACCCCCGCTGTGTGAAGCTGTGTACCACTGGAGAGAGCCCATCATCGCGCCGAG 1974
Db 1921 ACCCCCGCTGTGTGAAGCTGTGTACCACTGGAGAGAGCCCATCATCGCGCCGAG 1980
QY 1975 ACCTTCTACCTGGAGCGCGCCCAACCGCAGACCAAGATCGCAAGCGCGGTACTGTG 2034
Db 1981 ACCTTCTACCTGGAGCGCGCCCAACCGCAGACCAAGATCGCAAGCGCGGTACTGTG 2040
QY 2035 ACCGACCGGGCGCGCAGAGAAGATCGTGAAGCTGACCGAGACCAACCAAGAGACCGAG 2094
Db 2041 ACCGACCGGGCGCGCAGAGAAGATCGTGAAGCTGACCGAGACCAACCAAGAGACCGAG 2100
QY 2095 CTGACGGCCATCCAGCTGGCCCTGCAGGACAGCGGAGCGGTGAACATCGTGACCGAC 2154
Db 2101 CTGACGGCCATCCAGCTGGCCCTGCAGGACAGCGGAGCGGTGAACATCGTGACCGAC 2160
QY 2155 AGCAGTACCCCTGGGCATCATCCAGGCCAGCCGACAGAGCGAGAGCGAGCTGGTG 2214
Db 2161 AGCAGTACCCCTGGGCATCATCCAGGCCAGCCGACAGAGCGAGAGCGAGCTGGTG 2220
QY 2215 AACAGATCATCGAGCAGCTGATCAAGAGGAGAGTGTACCTGAGCTGGTGGCCGCC 2274
Db 2221 AACAGATCATCGAGCAGCTGATCAAGAGGAGAGTGTACCTGAGCTGGTGGCCGCC 2280
QY 2275 CACAAGGCGATCGCGCAAGCAGCAGATCGACAAGCTGTGAGCAAGGCGCATCCGCAAG 2334
Db 2281 CACAAGGCGATCGCGCAAGCAGCAGATCGACAAGCTGTGAGCAAGGCGCATCCGCAAG 2340
QY 2335 GTGCTGTCTCGACGGCATCGATGGCGCATCGTATCTACCAAGTACATGGAGACCTG 2394
Db 2341 GTGCTGTCTCGACGGCATCGATGGCGCATCGTATCTACCAAGTACATGGAGACCTG 2400
QY 2395 TAGTGGGCGAGCGCGCCCTAGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAA 2454
Db 2401 TAGTGGGCGAGCGCGCCCTAGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAA 2460
QY 2455 TTC 2457
Db 2461 TTC 2463

RESULT 3
ABL39959
ID ABL39959 standard; DNA; 2469 BP.
XX ABL39959;
XX AC
XX DT
XX 15-MAY-2002 (first entry)
XX Synthetic construct PR975(+) SEQ ID NO:30.
XX Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX Human immunodeficiency virus type C.
OS Synthetic.
XX WO200204493-A2.
XX PN
XX PD 17-JAN-2002.
XX PF 05-JUL-2001; 2001WO-US21241.
XX PR 05-JUL-2000; 2000US-0610313.

XX (CHIR) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX WPI; 2002-154920/20.
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
PT cell lines, particularly in gene therapy -
XX Claim 1; Fig 8; 233pp; English.
XX The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
CC Nef (1). (1) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the
CC exemplification of the present invention.
XX SQ Sequence 2469 BP; 571 A; 833 C; 761 G; 304 T; 0 other;
Query Match 98.3%; Score 2415.4; DB 24; Length 2469;
Best Local Similarity 99.3%; Pred. No. 2.8e-292;
Matches 2451; Conservative 0; Mismatches 5; Indels 12; Gaps 2;
QY 1 GTCGACGCCACCTGCGGAGGCCATGAGCCAGCCGACGCGCAACATCCTGATGCAG 60
DB 1 GTCGACGCCACCTGCGGAGGCCATGAGCCAGCCGACGCGCAACATCCTGATGCAG 60
QY 61 GCGAGCAACTTCAAGGCCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGAGGGC 120
DB 61 GCGAGCAACTTCAAGGCCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGAGGGC 120
QY 121 CACATCGCCGCAACTGCGGCGCCCGCCGCAAGAGGGCTGCTGAAGTGGCGCAAGAG 180
DB 121 CACATCGCCGCAACTGCGGCGCCCGCCGCAAGAGGGCTGCTGAAGTGGCGCAAGAG 180
QY 181 GGCCACCATGTAAGGACTGCAAGCGGCGGCGGCAAGTCTTCCGCGAGGACCTTGCC 240
DB 181 GGCCACCATGTAAGGACTGCAAGCGGCGGCGGCAAGTCTTCCGCGAGGACCTTGCC 240
QY 241 TTCCCGGAGGAGGCGCGGAGTTCCTCCAGCAGAGCAACCGCGCAACAGCCCCAAC 300
DB 241 TTCCCGGAGGAGGCGCGGAGTTCCTCCAGCAGAGCAACCGCGCAACAGCCCCAAC 300
QY 301 AGCGCGAGCTGCAAGTGGCGGCGCAACCCCGCAGCGAGGCGCGGCGGCGGCGGCG 360
DB 301 AGCGCGAGCTGCAAGTGGCGGCGCAACCCCGCAGCGAGGCGCGGCGGCGGCGGCG 360
QY 361 GGCACCTGAACTTCCCGGAGATCACCTGTGGCAGCGCCCGCTTGTGAGCATCAAGGTG 420
DB 361 GGCACCTGAACTTCCCGGAGATCACCTGTGGCAGCGCCCGCTTGTGAGCATCAAGGTG 420
QY 421 GCGCGGAGATCAAGGAGGCGCTGTCGACACCGCGCGGCGGCGGCGGCGGCGGCGG 480
DB 421 GCGCGGAGATCAAGGAGGCGCTGTCGACACCGCGCGGCGGCGGCGGCGGCGGCGG 480
QY 481 ATGAGCTGCGCGCAAGTGGAGGCCCAAGATGATCGCGGCGGCGGCGGCGGCGGCGG 540
DB 481 ATGAGCTGCGCGCAAGTGGAGGCCCAAGATGATCGCGGCGGCGGCGGCGGCGGCGG 540
QY 541 GTGCGGAGTACGACCATGCTGATCGAGATCTGCGGCGGCAAGAGGCGCATCGGCGCG 600
DB 541 GTGCGGAGTACGACCATGCTGATCGAGATCTGCGGCGGCAAGAGGCGCATCGGCGCG 600
QY 601 CTGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 660

DB 601 CTGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 660
QY 661 ACCCTGAACTTCCCGATCAGCCCATCGAGACCGTGGCGGAGAGTGAAGTGAAGCGGCG 720
DB 661 ACCCTGAACTTCCCGATCAGCCCATCGAGACCGTGGCGGAGAGTGAAGTGAAGCGGCG 720
QY 721 GACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
DB 721 GACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 780
QY 781 ATCTCGAGGAGATGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
DB 781 ATCTCGAGGAGATGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
QY 841 AACACCCCGGCTTCCCGATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 841 AACACCCCGGCTTCCCGATCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 901 TTCCGCGAGCTGAAACAGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
DB 901 TTCCGCGAGCTGAAACAGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
QY 961 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
DB 961 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
QY 1021 AGCGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
DB 1021 AGCGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
QY 1081 AACGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1081 AACGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
QY 1141 CCCAGCATCTTCCAGAGAGCATGACCAAGATCTTGGAGGCGGCGGCGGCGGCGGCGG 1200
DB 1141 CCCAGCATCTTCCAGAGAGCATGACCAAGATCTTGGAGGCGGCGGCGGCGGCGGCGG 1200
QY 1201 GAGATCGTGATCTACCA- - - - -GGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1254
DB 1201 GAGATCGTGATCTACCAAGATCTGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
QY 1255 CAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1314
DB 1255 CAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1314
QY 1315 CCCGACGAAGAGCAGCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1368
DB 1315 CCCGACGAAGAGCAGCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1368
QY 1369 GACAAGTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1428
DB 1369 GACAAGTGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1428
QY 1429 ATCCAGAGCTGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1488
DB 1429 ATCCAGAGCTGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1488
QY 1489 GCGCAGCTGTGCAAGTGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1548
DB 1489 GCGCAGCTGTGCAAGTGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1548
QY 1501 CGCCAGCTGTGCAAGTGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1560
DB 1501 CGCCAGCTGTGCAAGTGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1560
QY 1549 GAGGAGCGGAGCTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1608
DB 1549 GAGGAGCGGAGCTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1608
QY 1609 GTGTACTACCAAGTGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1668
DB 1609 GTGTACTACCAAGTGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1668
QY 1669 TGGACCTACCAAGTGTACCAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1728
DB 1669 TGGACCTACCAAGTGTACCAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1728

Db 541 CCGTGAAGCTGAAGCGGGGATGAGACGCCCAAGGTCAAGCACTGGCCCCCTGACCCGAGG 600
Qy 758 AGAAGATCAAGGCCCTTGACCGCCATCTCCGAGGAGATGAGAAAGAGGGCAAGATCACCA 817
Db 601 AGAAGATCAAGGCCCTTGATGAGATCTCCACCGAGATGGAGAAGAGGGCAAGATCACCA 660
Qy 818 AGATCGGCCCGGAGAACCCCTTACAAACACCCCGGTGTTCCGCCATCAAGAAGAAGACAGCA 877
Db 661 AGATCGGCCCGGAGAACCCCTTACAAACACCCCGGTGTTCCGCCATCAAGAAGAAGACAGCA 720
Qy 878 CCAAGTGCAGCAAGCTGTGTGACTTCCGCGAGCTGGAACAAGCGACACCCAGACTTCTGGG 937
Db 721 CCAAGTGCAGCAAGCTGTGTGACTTCCGCGAGCTGGAACAAGCGACACCCAGACTTCTGGG 780
Qy 938 AGGTGCAGCTGGGATCCCCACCCCGCGGCTGGAAGAAGAAGAGCGTGACCGTGC 997
Db 781 AGGTGCAGCTGGGATCCCCACCCCGCGGCTGGAAGAAGAAGAGCGTGACCGTGC 840
Qy 998 TGGAGCTGGGCGAGCGCTTACTTACGCGTGCCTGGACAAGGACTTCCGCAAGTACACCG 1057
Db 841 TGGAGCTGGGCGAGCGCTTACTTACGCGTGCCTGGACAAGGACTTCCGCAAGTACACCG 900
Qy 1058 CCTTCACCATCCCGAGCATCAACACGAGACCCCGGCGATCCGCTACCAAGTACAACTGC 1117
Db 901 CCTTCACCATCCCGAGCATCAACACGAGACCCCGGCGATCCGCTACCAAGTACAACTGC 960
Qy 1118 TGCCTCCAGGGCTGGAAGGGCAAGCCCGACATCTTCCAGAGCAGCATGACCAAGATCTCTGG 1177
Db 961 TGCCTCCAGGGCTGGAAGGGCAAGCCCGGCGATCTTCCAGAGCAGCATGACCAAGATCTCTGG 1020
Qy 1178 AGCCCTTCCGCGCGCGCAACCCCGAGATCGTGATCTACCAAGGCCCGCTGTACGTGGSCA 1237
Db 1021 AGCCCTTCCGCGCGCGCAACCCCGAGATCGTGATCTACCAAGGCCCGCTGTACGTGGSCA 1080
Qy 1238 GCGACCTGGAGATCGCCAGCAGCCCGCCAGATCGAGGAGCTCGGCAAGCACCTGCTGC 1297
Db 1081 GCGACCTGGAGATCGCCAGCAGCCCGCCAGATCGAGGAGCTCGGCGAGCACCTGCTGC 1140
Qy 1298 GCTGGGGCTTCCACACCCCGCACAAAGACACCAAGAGAGCCCGCTTCTGCGCCATCG 1357
Db 1141 GCTGGGGCTTCCACACCCCGCACAAAGACACCAAGAGAGCCCGCTTCTGCGCCATCG 1200
Qy 1358 AGCTGCACCCCGACAAGTGGACCTGACGCCATCGAGCTCGCCGAGAGAGAGCTGGA 1417
Db 1201 AGCTGCACCCCGACAAGTGGACCTGACGCCATCGAGCTCGCCGAGAGAGAGCTGGA 1260
Qy 1418 CCCTGAACGACATCCAGAAGCTGTGGCAAGCTGMACTGGCGCAGCCAGATCTACCCCG 1477
Db 1261 CCCTGAACGACATCCAGAAGCTGTGGCAAGCTGMACTGGCGCAGCCAGATCTACCCCG 1320
Qy 1478 GCATCAAGGTGCGCCAGCTGTGCAAGCTGTGCGCGCGCCCAAGGCCCTGACCGACATCG 1537
Db 1321 GCATCAAGGTGCGCCAGCTGTGCAAGCTGTGCGCGCGCCCAAGGCCCTGACCGAGTGA 1380
Qy 1538 TGCCCTTGACCGAGAGCGGAGCTGGAGCTGGCCGAGAACCGGAGATCTGCGCGAGC 1597
Db 1381 TGCCCTTGACCGAGAGCGGAGCTGGAGCTGGCCGAGAACCGGAGATCTGCGCGAGC 1440
Qy 1598 CCCTGACGCGGTGCTACTAGACCCCGACGAGGACCTGGTGGCGGAGATCCAGAGCAGG 1657
Db 1441 CCCTGACGCGGTGCTACTAGACCCCGACGAGGACCTGGTGGCGGAGATCCAGAGCAGG 1500
Qy 1658 GCCACGACAGTGGACCTTACAGATCTACCAAGAGCCCTTCAAGAACTGAAGACCGGCA 1717
Db 1501 GCCAGGCGAGTGACCTTACAGATCTACCAAGAGCCCTTCAAGAACTGAGACCGGCA 1560
Qy 1718 AGTACGCCAAGATCGGCACCGCCCGCACACACGACGTGAAGAGCTGACCGAGCCGTGC 1777
Db 1561 AGTACGCCCGCATGCGCGGGCGCCACACCAAGACGTGAAGCAGCTGACCGAGCCGTGC 1620
Qy 1778 AGAAGATCGGCATGAGAGCATCTGTATCTGGGCAAGACCCCAAGTTCCGCTGCGCA 1837
Db 1621 AGAAGTGAACCGAGAGCATCTGTATCTGGGCAAGATCTCCCAAGTTCAAGCTGCCCA 1680

Qy 1838 TCCAGAAGGAGACCTGGGAGACCTGCTGACCGCACTACTGGCAGGCCACCTGGATCCCCG 1897
Db 1681 TCCAGAAGGAGACCTGGGAGGCTTGTGATGGAGTACTTGGCAGGCCACCTGGATCCCCG 1740
Qy 1898 AGTGGAGTTCGTGAACACACCCCGCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGC 1957
Db 1741 AGTGGAGTTCGTGAACACACCCCGCTGGTGAAGCTGTGGTACCAGCTGGAGAAGGAGC 1800
Qy 1958 CCATCATCGGCGCCGAGACCTTCTACGTGGAGCGCGCCCAACCGCGAGACCAAGATCG 2017
Db 1801 CCATCGTGGCGCCGAGACCTTCTACGTGGAGCGCGCCCAACCGCGAGACCAAGCTGG 1860
Qy 2018 GCAAGCGCGGTACGTGACCGACCGCGCGCAGAACTGTGAGCTTGACCTGACCGAGACCA 2077
Db 1861 GCAAGCGCGGTACGTGACCGACCGCGCGCGCAGAACTGTGAGCTTGACCTGACCGAGACCA 1920
Qy 2078 CCAACAGAAGACCGAGCTGACGAGGCATCCAGCTGGCCCTGCAGGACAGCGCAGCGAGG 2137
Db 1921 CCAACAGAAGACCGAGCTGACGAGGCATCCAGCTGGCCCTGCAGGACAGCGCAGCGAGG 1980
Qy 2138 TGAACATCGTGACCGACAGCCAGTACGCCCTGGGCGATCATCCAGGCCAGCCCGACAAGA 2197
Db 1981 TGAACATCGTGACCGACAGCCAGTACGCCCTGGGCGATCATCCAGGCCAGCCCGACAAGA 2040
Qy 2198 GCGAGAGCAGCTGTGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACC 2257
Db 2041 GCGAGAGCAGCTGTGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACC 2100
Qy 2258 TGAGTGTGGTGCCCGCCACAAAGGCATCGCGCGCAAGCAGCAGATCGACAAAGCTGGTGA 2317
Db 2101 TGCGCTGGTGCCCGCCACAAAGGCATCGCGCGCAAGCAGCAGCTGGACAAAGCTGGTGA 2160
Qy 2318 GCAAGGGCATCCGCAAGTGTCTTCTGGAGCGCATCGATGGCGCATCTGATCTTACC 2377
Db 2161 GCGCGGCGATCCGCAAGTGTCTTCTTGAAGCGCATCGATGGCGCATCTGATCTTACC 2220
Qy 2378 AGTACATGACGACCTGTACGTGGGCGCGCGCCCTAGGATCGATTAAAGCTTCCCCG 2437
Db 2221 AGTACATGACGACCTGTACGTGGGCGCGCGCCCTAGGATCGATTAAAGCTTCCCCG 2380
Qy 2438 GGGCTAGCACCGGTGAATTC 2457
Db 2281 GGGCTAGCACCGGTGAATTC 2300

RESULT 5

AAA70479

ID AAA70479 standard; DNA; 2306 BP.

AC AAA70479;

XX

DT 28-NOV-2000 (first entry)

XX HIV FS(-)_ProtMod_rTOpt_YM coding sequence.

DE HIV-1; AIDS; Gag-reverse transcriptase; vaccine; expression cassette; ss.

XX Human immunodeficiency virus type 1.

OS Synthetic.

XX WO200039302-A2.

XX 06-JUL-2000.

PD

XX 30-DEC-1999; 99WO-US31245.

PF 31-DEC-1998; 98US-0114495.

PR 01-DEC-1999; 99US-0168471.

XX (CHIR) CHIRON CORP.

XX Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;

PI Greer C, Selby M, Walker C;

XX WPI; 2000-452400/39.

XX Expression cassettes encoding the human immunodeficiency virus (HIV)

PT Gag-containing polypeptide useful for vaccinating against HIV

PT Infections and acquired immunodeficiency syndrome (AIDS) -

XX Claim 7; Fig 73; 391pp; English.

XX The present sequence is the coding sequence of a synthetic HIV

CC Gag-reverse transcriptase expression cassette, FS(-)_ProtMod_Rtopt_YM.

CC The Gag protein of HIV is needed for the assembly of virus-like

CC particles. In addition, the Gag protein is involved in many stages of the

CC HIV life cycle, including assembly, virion maturation after particle

CC release and early post-entry steps in viral replication. The expression

CC cassette may be used for the recombinant expression of HIV

CC Gag-polypeptides which may then be used to vaccinate against HIV

CC infection and acquired immunodeficiency syndrome (AIDS).

XX Sequence 2306 BP; 529 A; 752 C; 730 G; 295 T; 0 other;

Query Match 82.2%; Score 2019.2; DB 21; Length 2306;

Best Local Similarity 93.2%; Pred. No. 4.8e-243;

Matches 2150; Conservative 0; Mismatches 138; Indels 18; Gaps 3;

QY 170 GCGCAAGAGGAGGCGCACAGATGAAGACTGCACCGAGCGCCAGGCCAACTTCTTCGCG 229

DB 1 GCGGCGCGAAGACACCAATGAAGATTGCACTGAGACAGGCTAATTCTTCGCG 60

QY 230 AGGACCTGGCTTCCGCCAGGCGAAGCCGCGGAGTTCCCGCAGCAGCAACCGGCCA 289

DB 61 AGGACCTGGCTTCCCGCAGGCGAAGCCGCGGAGTTCCAGCAGCAGCAGCCGCGCCA 120

QY 290 ACAGCCCCACAGCCGAGCTGCAAGTGGCGGCGG-----ACAACCCCGCAGCGAGG 343

DB 121 ACAGCCCCACCGCGGAGCTGCAGTGTGGGCGCGAGAACACAGCCTGAGCGAGG 180

QY 344 CCGCGCGCAGGCGCAGGCGACCTG-----AACTTCCCGCAGATCACCTGTGGCAGC 397

DB 181 CCGGCGCGACCGCAGGCGACCTGAGCTTCACTTCCCGCAGATCACCTGTGGCAGC 240

QY 398 GCGCCCTGGTGAGCATCAAGTGGCGGCGCGAGATCAAGGAGGCGCTCTGCGACCGCGG 457

DB 241 GCGCCCTGGTGAGCATCAGGATCGCGCGCAGCTCAAGGAGGCGCTCTGCGACCGCGG 300

QY 458 CCGACGACACCGTCTGGAGGAGATGAGCTGCCCCGCAAGTGGAAAGCCCAAGATGATCG 517

DB 301 CCGACGACACCGTCTGGAGGAGATGAACCTGCGCGGCAAGTGGAAAGCCCAAGATGATCG 360

QY 518 GCGCATCGCGGCTTCATCAAGTGGCGCGCAGTACGACACAGATCCTGATCGAGATCTGCG 577

DB 361 CCGGATCGGGGCTTCATCAAGTGGCGCAGTACGACACAGATCCTCGTGAGATCTGCG 420

QY 578 GCAAGAGGCCATCGGCACCGTGTGATCGGCGCGCCACCGCGTGAACATCATCGGCGCA 637

DB 421 GCCACAAGGCCATCGGCACCGTGTGTTGGGCGCGCCACCGCGTGAACATCATCGGCGCA 480

QY 638 ACATGCTGACCCAGCTGGGCTGACCTGAACTTCCCGCATCGCCCATCGAGACCGTGC 697

DB 481 ACCTGCTGACCCAGATCGGCTGACCTGAACTTCCCGCATCGCCCATCGAGACCGTGC 540

QY 698 CCGTGAAGCTGAAGCCCGGATGACGCGCGCCCAAGTGAAGCAGTGGCGCTGACCGAGG 757

DB 541 CCGTGAAGCTGAAGCCCGGATGACGCGCGCCCAAGTGAAGCAGTGGCGCTGACCGAGG 600

QY 758 AGAAGATCAAGGCCCTGACCGCATCTGCGAGGAGATGGAGAGGAGGGAAGATCACCA 817

DB 601 AGAAGATCAAGGCCCTGCTGGAGATCTGCACGAGATGGAGAGGAGGGAAGATCACCA 660

QY 818 AGATCGCGCGGAGACCCCTACACACCCCGTGTTCGCCATCAAGAAGAGGACAGCA 877

DB 661 AGATCGCGCGGAGACCCCTACACACCCCGTGTTCGCCATCAAGAAGAGGACAGCA 720

QY 878 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG 937
DB 721 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGG 780
QY 938 AGTGCAGCTGGGATCCGCCACCCCGCGCTGGAAGAAGAGAGAGGCTGACCTGCG 997
DB 781 AGTGCAGCTGGGATCCGCCACCCCGCGCTGGAAGAAGAGAGAGGCTGACCTGCG 840
QY 998 TGGACGTGGGCGAGCGCTACTTCAGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCG 1057
DB 841 TGGACGTGGGCGAGCGCTACTTCAGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCG 900
QY 1058 CTTTACCATCTCCCGAGCATCAACAACAGAGACCCCGCGCATCCCGTACAGTACACCTGC 1117
DB 901 CTTTACCATCTCCCGAGCATCAACAACAGAGACCCCGCGCATCCCGTACAGTACACCTGC 960
QY 1118 TGGCCCCAGGCTGAAGGGGAGGCGCCAGCATCTTCCAGAGAGGATGACCAAGATCCTGG 1177
DB 961 TGGCCCCAGGCTGAAGGGGAGGCGCCCGCATCTTCCAGAGAGGATGACCAAGATCCTGG 1020
QY 1178 AGCCCTTCCCGCGCGCAACCCCGAGATCGTGTATCTACAGGCCCGCTGTACCTGGGCA 1237
DB 1021 AGCCCTTCCCGCAAGAGACCCCGCATCTGTATCTACAGGCCCGCTGTACCTGGGCA 1080
QY 1238 GCGACCTTGGAGATCGGCGCAGCACCGCCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGC 1297
DB 1081 GCGACCTTGGAGATCGGCGCAGCACCGCCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGC 1140
QY 1298 GCTGGGGCTTACCAACCCCGCAGCAAGAACACCAAGAGGAGGCGCCCTTCTGCGCCAT-- 1355
DB 1141 GCTGGGGCTTACCAACCCCGCAGCAAGAACACCAAGAGGAGGCGCCCTTCTGCTGGATGG 1200
QY 1356 ----CGAGCTGCACCCCGCAAGTGGACCGTGCAGCCCATCGAGCTGCGCGAGAGGAGA 1411
DB 1201 GCTACGAGCTGCACCCCGCAAGTGGACCGTGCAGCCCATCATCTGCTGCCGAGAGGACA 1260
QY 1412 GCTGGACCGTGAAGCAGATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGCCAGCCAGATCT 1471
DB 1261 GCTGGACCGTGAAGCAGATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGCCAGCCAGATCT 1320
QY 1472 ACCCGCGCATCAAGTGGCGCAGCTGTGCAAGCTGTGCGGGGCGCAAGGCCCTGACCG 1531
DB 1321 ACGCCGCGCATCAAGCTGAAGCAGCTGTGCAAGCTGTGCGGGGCGCAAGGCCCTGACCG 1380
QY 1532 ACATCGTGCCTTGCACCGAGGCGCCAGCTGGAGCTGGCGGAGAACCGCGAGATCTTGC 1591
DB 1381 AGTGTATCCCCCTGACCGAGGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTTGA 1440
QY 1592 GCGAGCCCGTGCACGGCGTGTACTACGACCCCGCAGCAAGGACCTGGTGGCGGAGATCCAGA 1651
DB 1441 AGGAGCCCGTGCACGGCGTGTACTACGACCCCGCAGCAAGGACCTGGTGGCGGAGATCCAGA 1500
QY 1652 AGCAGGCCAGGACCGAGTGGACCTTACAGATCTTACAGGAGGCCCTTCAGAACCTGAAGA 1711
DB 1501 AGCAGGCCAGGCGGCGAGTGGACCTTACAGATCTTACAGGAGGCCCTTCAGAACCTGAAGA 1560
QY 1712 CCGCAAGTACGCCAAGATGCGCACCGCCCGCAGCAACACCAAGCTGAAGCAGCTGACCGAGG 1771
DB 1561 CCGCAAGTACGCCCGCATGCGCGGCGCGCCACACCAAGCTGAAGCAGCTGACCGAGG 1620
QY 1772 CCGTGAAGAAGATCGCGCATGGAGCATCTGTATCTTGGGCGAAGACCCCGCAAGTTCGCGC 1831
DB 1621 CCGTGAAGAAGTGGACCGAGAGCATCTGTATCTTGGGCGAAGATCCCCAAGTTCGAAGC 1680
QY 1832 TGGCCATCCAGAAGGAGACCTGGAGACCTTGGTGGACCGGACTACTTGGCAGGCCACCTTGA 1891
DB 1681 TGGCCATCCAGAAGGAGACCTGGAGACCTTGGTGGATCTGTGCGAGGCCACCTTGA 1740
QY 1892 TCCCGGAGTGGGAGTCTGTGAACACCCCGCTGGTGAAGCTGTGGTACCAAGCTGGAGA 1951
DB 1741 TCCCGGAGTGGGAGTCTGTGAACACCCCGCTGGTGAAGCTGTGTACCAAGCTGGAGA 1800

XX The present sequence is the coding sequence of a synthetic HIV
CC gag-reverse transcriptase expression cassette, FS(-).ProtMod.Ropt(+).
CC The Gag protein of HIV is needed for the assembly of virus-like
CC particles. In addition, the Gag protein is involved in many stages of the
CC HIV life cycle, including assembly, virion maturation after particle
CC release and early post-entry steps in viral replication. The expression
CC cassette may be used for the recombinant expression of HIV
CC gag-polypeptides which may then be used to vaccinate against HIV
CC infection and acquired immunodeficiency syndrome (AIDS).
XX
SQ Sequence 2312 BP; 533 A; 750 C; 732 G; 297 T; 0 other;

Query Match 81.3%; Score 1998.4; DB 21; Length 2312;
Best Local Similarity 92.9%; Pred. No. 1.8e-240;
Matches 2147; Conservative 0; Mismatches 141; Indels 24; Gaps 4;

QY 170 GCGCAAGGAGGCGCCACCATGAAAGACTGCACCGAGCGCGCCACCACTTCTTCGCG 229
DB 1 GCGGCGCGAAGGACACCAATGAAGATTGCACTGAGAGACAGGCTAATTTCTTCGCG 60
QY 230 AGGACCTGGCTTCCCGCAGGCAAGCGCGAGTTCCCGAGCGAGCAAGCGCGCA 289
DB 61 AGGACCTGGCTTCTGAGGCGAAGCGCGAGTTGAGAGCGAGACCGCGCA 120
QY 290 ACAGCCCCACAGCGCGGAGCTGCAGGTGCGCGCGG-----ACAACCCCCCGAGCGAGG 343
DB 121 ACAGCCCCACCGCGCGAGCTGCAGGTGCGCGCGGAGAAACACAGCTGAGCGAGG 180
QY 344 CCGCGCGGAGCGCGACCGACCTG-----AACTTCCCCCAGATCACCTGTGGGAGC 397
DB 181 CCGCGCGGAGCGCGACCGGACCGTGAAGTTCAACTTCCCCCAGATCACCTGTGGGAGC 240
QY 398 GCCCCTGGTGAGCATCAAGTGGCGCGCAGATCAAGAGGCGCTGCTGCACACCGCG 457
DB 241 GCCCCTGGTGACCATCAGATCGCGCGCAGCTCAAGAGGCGGTGCTGCACACCGCG 300
QY 458 CCGAGGACACCGTGTGAGGAGATGAGCTGCGCGGCAAGTGAAGCCCCAAGATGATCG 517
DB 301 CCGAGGACACCGTGTGAGGAGATGAGCTGCGCGGCAAGTGAAGCCCCAAGATGATCG 360
QY 518 CCGCATCGCGCGCTTCATCAAGTGGCGCAGTACGACGAGCGCTGCTGCACACCGCG 577
DB 361 CCGGATCGCGCGCTTCATCAAGTGGCGCAGTACGACGAGCTCCCGTGAGATCTGCG 420
QY 578 GCAAGAGGCGCATCGGACCGCTGTGATCGCGCCACACCGCGTGAACATCATCGCGCGCA 637
DB 421 GCCAAGGCGCATCGGACCGCTGTGTTGGCGCCACCGCGTGAACATCATCGCGCGCA 480
QY 638 ACATGCTGACCGCTGGGTGACCTGTGAATCTCCCGCATCAGCCCGCATCGAGACCGTGC 697
DB 481 ACCTGCTGACCGCATCGGCTGCACTTGAATCTCCCGCATCAGCCCGCATCGAGCGGTGC 540
QY 698 CCGTGAAGCTGAAGCGCGCATGACGCGCCCAAGGTGAAGCTGAGCGCTGACCGAGG 757
DB 541 CCGTGAAGCTGAAGCGCGCATGACGCGCCCAAGGTGAAGCTGAGCGCTGACCGAGG 600
QY 758 AGAAGATCAAGGCGCTGACCGCATCTCGGAGGAGATGAGAGGAGGCGCAAGATCACCA 817
DB 601 AGAAGATCAAGGCGCTGAGAGATCTGCACCGAGATGAGAGGAGGCGCAAGATCACCA 660
QY 818 AGATCGCGCGGAGAACCCCTTACAAACCGCCGTGTTGCGCATCAAGAAGAGGACAGCA 877
DB 661 AGATCGCGCGGAGAACCCCTTACAAACCGCCGTGTTGCGCATCAAGAAGAGGACAGCA 720
QY 878 CCAAGTGGCGCAAGCTGTGAGTTCGCGGAGTGAACACCGCGTGTGCGCATCAAGAAGAGGACAGCA 937
DB 721 CCAAGTGGCGCAAGCTGTGAGTTCGCGGAGTGAACACCGCGTGTGCGCATCAAGAAGAGGACAGCA 780
QY 938 AGGTGACGCTGGGATCCCCACCGCGCGCTGGAAGAGAGAGAGGCTGACCGTGC 997
DB 781 AGGTGACGCTGGGATCCCCACCGCGCGCTGGAAGAGAGAGAGGCTGACCGTGC 840

QY 998 TGGAGCTGGCGAGCGCTACTTTCAGCGTCCCTGAGCAGGAGCTTCCGCAAGTACACG 1057
DB 841 TGGAGCTGGCGAGCGCTACTTTCAGCGTCCCTGAGCAGGAGCTTCCGCAAGTACACG 900
QY 1058 CTTTCAACATCCCGAGCTCAACAAGAGAGACCGCGCATCGGCTACCAAGTCAACAGTGC 1117
DB 901 CTTTCAACATCCCGAGCTCAACAAGAGAGACCGCGCATCGGCTACCAAGTCAACAGTGC 960
QY 1118 TGCCCCAGGGCTGGAAGGCGACCGCGCATCTTCAGAGCAGCATGACCAAGATCCTGG 1177
DB 961 TGCCCCAGGGCTGGAAGGCGACCGCGCATCTTCAGAGCAGCATGACCAAGATCCTGG 1020
QY 1178 AGCCCTTCCGCGCGCAACCGCGAGATCGTGTATCTACCA-----GGCCCCCTGTAGC 1231
DB 1021 AGCCCTTCCGCGCGCAACCGCGAGATCGTGTATCTACCAAGTACATGAGCAGCTGTAGC 1080
QY 1232 TGGGAGGAGACCTGAGATCGGCCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCACC 1291
DB 1081 TGGGAGGAGACCTGAGATCGGCCAGCACCGCGCAAGATCGAGGAGCTGCGCCAGCACC 1140
QY 1292 TGCTCGGCTGGGCTTCACCAACCGCGCGACAAAGACCAAGAGAGCGCCCTTCTCTGC 1351
DB 1141 TGCTCGGCTGGGCTTCACCAACCGCGCGACAAAGACCAAGAGAGCGCCCTTCTCTGC 1200
QY 1352 CCAT-----CGAGCTGCACCGCGACAAAGTGCAGCGCTGCAGCGCATCGAGCTGCCCGAGA 1405
DB 1201 GGATGGGCTACGAGCTGCACCGCGACAAAGTGCAGCGCTGCAGCGCATCTGCTGCCCGAGA 1260
QY 1406 AGGAGAGTGGACCTGAAACGACATCCAGAACTGGTGGGCAAGTGAATCGGGCAGCG 1465
DB 1261 AGGAGAGTGGACCTGAAACGACATCCAGAACTGGTGGGCAAGTGAATCGGGCAGCG 1320
QY 1466 AGATCTACCGCGCATCAAGTGGCGCGAGCTGTGCAAGCTGCTGCGCGCGCGCAAGCGCC 1525
DB 1321 AGATCTACCGCGCATCAAGTGAAGCTGTGCAAGCTGCTGCGCGCGCGCAAGCGCC 1380
QY 1526 TGACCGACATCGTGCCTCCCTGACCGAGGAGCGAGCTGGAGCTGCGCGAAGACCGCGAGA 1585
DB 1381 TGACCGAGTGTATCCCTGACCGAGGAGCGAGCTGGAGCTGCGCGGAGAACCGCGAGA 1440
QY 1586 TCCTGCGGAGCGCTGACCGCGTGTACTAGACCGCAGCAAGACCTGTGTTGGCGCGAGA 1645
DB 1441 TCCTGAGAGGAGCGCTGACCGAGTGTACTAGACCGCAGCAAGACCTGTGTTGGCGCGAGA 1500
QY 1646 TCCAGAAAGCGGCGCACCGACAGTGCAGCTACCAAGTGTGAGGAGCGCTTTCAGAAACC 1705
DB 1501 TCCAGAAAGCGGCGCGAGCGAGTGGAGCTTACAGATCTTACCGAGGCGCTTTCAGAAACC 1560
QY 1706 TGAAGACCGGCAAGTACCGCAAGATGCGCACCGCGCCACCAACACGCTGAAGCAGCTGA 1765
DB 1561 TGAAGACCGGCAAGTACCGCGCATGCGCGGCGCGCCACACCAACACGCTGAAGCAGCTGA 1620
QY 1766 CCGAGGCGCTGCAGAAAGATCGCATGAGAGAGATCGTGTATGTTGGGCGCAAGACCCCAAGT 1825
DB 1621 CCGAGGCGCTGCAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1680
QY 1826 TCCGCTTCCCATCCAGAAAGAGAGCTGCGAGAGCTGTTGAGAGCTGTTGAGAGCTGTTGAG 1885
DB 1681 TCAAGCTGCCCATCCAGAAAGAGAGCTTGGAGGCGCTTGGTGAAGCTGTTGAGAGCTGTTGAG 1740
QY 1886 CCTGGATCCCGAGTGGAGTTCGTGAACACCGCGCGCTGTTGAGAGCTGTTGAGAGCTGTTGAG 1945
DB 1741 CCTGGATCCCGAGTGGAGTTCGTGAACACCGCGCGCTGTTGAGAGCTGTTGAGAGCTGTTGAG 1800
QY 1946 TGGAGAGGAGCGCATCATCGCGCGCGAGAGCTTCTTACGTGAGAGCGCGCGCGAGAGATCGTGAGCG 2005
DB 1801 TGGAGAGGAGCGCATCATCGCGCGCGAGAGCTTCTTACGTGAGAGCGCGCGCGAGAGATCGTGAGCG 1860
QY 2006 AGACCAAGATCGGCAAGCGCGCTACGTGAGCGCGCGCGCGCGAGAGATCGTGAGCG 2065
DB 1861 AGACCAAGATCGGCAAGCGCGCTACGTGAGCGCGCGCGCGAGAGATCGTGAGCG 1920
QY 2066 TGACCGAGACCGCAACAGCAAGAGACCGAGCTGCGAGCGCATCCAGCTGCGCGCTGCGAGGACA 2125

Db 1921 TCAGCGACACCAACCAAGAGCGAGCTGCAGGCGCATCCACCTGGCCCTGCAGSACA 1980
 QY 2126 GCGGACGAGAGTGACATCGTGACGACAGCAGAGTACGCGCTGGGCTATCATCAGGCC 2185
 Db 1981 GCGGCTGGAGGTGAACATCGTGACCGACAGCAGTACGCGCTGGGCTATCATCAGGCC 2040
 QY 2186 AGCCGCAAGAGCGAGGAGCTGGTGAACAGCATCATCGAGCAGCTGATCAAGAAGG 2245
 Db 2041 AGCCGCAAGAGCGAGGAGCGAGCTGGTGAACAGCATCATCGAGCAGCTGATCAAGAAGG 2100
 QY 2246 AGAAGGTGTACCTGAGCTGGTGGCCGCCCAAGAGGSCATCGCGGCAACGACAGATCG 2305
 Db 2101 AGAAGGTGTACCTGAGCTGGTGGCCGCCCAAGAGGSCATCGCGGCAACGACAGTGG 2160
 QY 2306 ACAAGCTGGTGAGCAAGGCGATCGGCAAGTGTCTTGGACGGCATCGATGGCGGCA 2365
 Db 2161 ACAAGCTGGTGAGCGCGGCGATCGGCAAGTGTCTTGGACGGCATCGATGGCGGCA 2220
 QY 2366 TCGTGATCTACCACTACATGAGCAGCTGTACGTGGGCGGCGGCTTAGGATCGATT 2425
 Db 2221 TCGTGATCTACCACTACATGAGCAGCTGTACGTGGGCGGCGGCTTAGGATCGATT 2280
 QY 2426 AAAAGCTTCCCGGGCTAGCACCGGTGAATTC 2457
 Db 2281 AAAAGCTTCCCGGGCTAGCACCGGTGAATTC 2312

RESULT 8

ABK91622

ID ABK91622 standard; DNA; 9788 BP.

AC ABK91622;

XX 14-AUG-2002 (first entry)

DE Modified HIV protein-encoding plasmid DNA #174.

XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
 KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
 KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
 XX Synthetic.

WO200232943-A2.

XX 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.

XX 14-AUG-2000; 2000US-225097P.

XX 14-NOV-2000; 2000US-252115P.

XX 28-MAR-2001; 2001US-279257P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX (CHAD/) CHADABARTI B K.

XX Nabel GJ, Huang Y;

XX NPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and

XX its encoded protein, useful as vaccines for genetic or protein

XX immunisation for acquired immunodeficiency syndrome or human

XX immunodeficiency virus infection

XX Disclosure; Page 785-788; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV

XX (human immunodeficiency virus) protein. The nucleic acid molecule and its

XX encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for

XX genetic or protein immunisation to a host, respectively. In particular

XX these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
 CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
 CC compositions are useful for treating or preventing HIV infections or
 CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
 CC plasmid DNA sequences of the invention.
 XX

SQ Sequence 9788 BP; 2377 A; 2817 C; 2695 G; 1899 T; 0 other;

Query Match 81.1%; Score 1992.4; DB 24; Length 9788;

Best Local Similarity 91.2%; Pred. No. 8.4e-240;

Matches 2181; Conservative 0; Mismatches 181; Indels 30; Gaps 5;

QY 14 TGGCCGAGGCGCATGAGCGAGGCCACCAACATCTCTGATCGAGCGCAACTTCA 73

Db 2561 TGGCCGAGGCGCATGAGCGAGGTGAACAACCAACATCATGATCGAGCGCAACTTCA 3020

QY 74 AGGCGCCCAAGCGCATCATCAAGTGTCTTAAGTCTGCGGCAAGGAGGCCACATCCCGCGCA 133

Db 3021 AGGCGCCCAAGCGCATCAAGTGTCTTAAGTCTGCGGCAAGGAGGCCACATCCCGCGCA 3080

QY 134 ACTGCCGCGCCCGCGCAAGAGGCTGTGGAAGTGCAGGAGGAGGCCACCAAGTGA 193

Db 3081 ACTGCCGCGCCCGCGCAAGAGGCTGTGGAAGTGCAGGAGGAGGCCACCAAGTGA 3140

QY 194 AGGACTGCACCGAGCGCGCCCAACTTCTTCCGCGAGGACCTTGGCCCTTCCCGCAGGGCA 253

Db 3141 AGGACTGCACCGAGCGCGCCCAACTTCTTCCGCGAGGACCTTGGCCCTTCCCGCAGGGCA 3194

QY 254 AGGCGCGCGAGTTCGCCAGGAGCAGAGAACCGCGCAACAGCCCGCCACAGCGCGCGAGTGC 313

Db 3195 AGGCGCGCGAGTTCGCCAGGAGCAGAGAACCGCGCAACAGCCCGCCACAGCGCGAGTGC 3254

QY 314 AGGTGCGCGG-----CGACAACCCCGCGAGGCGCGCGCGAGCGCGCGCGCGCGCA--- 364

Db 3255 AGGTGCGCGGAGAGACAACAACTCCCTCTCAGAGGAGGAGCGCGCGCGAGCGCGCGCGCGCA 3314

QY 365 ---CCCTGAATCTCCCGCAGATCACTCCCTGTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 421

Db 3315 TATCCTTTAGTCTCCCTCAGATCACTCTTTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 3374

QY 422 GCGCGCAGATCAAGGAGGCGCTGTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 481

Db 3375 GGGCGCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3434

QY 482 TGAGCCTGCCCGCAAGTGGAAAGCCCAAGATGATCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 541

Db 3435 TGAACCTGCCCG 3494

QY 542 TGCGCCAGTACGACACAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCGCGCGCGCG 601

Db 3495 TGGGCCAGTACGACACAGATCCTGATCGAGATCTGCGGCGCGCGCGCGCGCGCGCGCGCGCG 3554

QY 602 TGATCG 661

Db 3555 TGGTGGCG 3614

QY 662 CCCTGAATCTCCCGCAGATCCTGATCGAGATCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 721

Db 3615 CCCTGAATCTCCCGCAGATCCTGATCGAGATCTGCGGCGCGCGCGCGCGCGCGCGCGCGCG 3674

QY 722 ACGCCCG 781

Db 3675 ACGCCCG 3734

QY 782 TCTGCGAGGAGATGGAGAGGAGGCGCAAGATCAACAGATCGCGCGCGCGCGCGCGCGCGCGCG 841

Db 3735 TCTGCGAGGAGATGGAGAGGAGGCGCAAGATCAACAGATCGCGCGCGCGCGCGCGCGCGCGCG 3794

QY 842 ACACCG 901

Db 3795 ACACCG 3854

QY	902	TCCGCGAGCTGAACAAGCGACCCAGGACTTCTGGAGGTGCAGCTGGGCAATCCCCCAACC	961
DB	3855	TCCGCGAGCTGAACAAGCGACCCAGGACTTCTGGAGGTGCAGCTGGGCAATCCCCCAACC	3914
QY	962	CCGCGCGCTGAAGAAGAAAGAGCGTGACCGTGTGTGACGTGGCGGCAACGCTACTTCA	1021
DB	3915	CCGCGCGCTGAAGAAGAAAGAGCGTGACCGTGTGTGACGTGGCGGCAACGCTACTTCA	3974
QY	1022	CGGTGCCCTTGGAGGAGACTTCCGCAAGTACACGCTTCCACCTCCCAAGCATCAACA	1081
DB	3975	CGGTGCCCTTGGAGGAGACTTCCGCAAGTACACGCTTCCACCTCCCAAGCATCAACA	4034
QY	1082	ACGAGACCCCGGCATCCGCTACCACTACAAAGTGTGCTGCCCGGAGGCTGGAAGGCAAGC	1141
DB	4035	ACGAGACCCCGGCATCCGCTACCACTACAAAGTGTGCTGCCCGGAGGCTGGAAGGCAAGC	4094
QY	1142	CCAGCATCTTCCAGAGCAGCATGACCACAGATCCTTGGAGCCTTCCGGCCCGCAACCCCG	1201
DB	4095	CCGCCATCTTCCAGTGCAGCATGACCACAGATCCTTGGAGCCTTCCGGCCCGCAACCCCG	4154
QY	1202	AGATCGTGATCTACCA-----GGCCCCCTGTACGTGGGAGCGACCTGGAGATCGGCC	1255
DB	4155	ACATCGTGATCTACCAAGTACATGGACCACCTGTACGTGGGAGCGACCTGGAGATCGGCC	4214
QY	1256	AGCACCGCGCAAGATCGAGGAGCTGCGCAACGACCTGCTGGGCTGGGCTTACACACC	1315
DB	4215	AGCACCGCGCAAGATCGAGGAGCTGCGCAACGACCTGCTGGGCTGGGCTTACACACC	4274
QY	1316	CCGACAAAGAACCAAGAGGAGCCCTCTCTGCCCAT-----CGAGCTGCACCCCG	1369
DB	4275	CCGACAAAGAACCAAGAGGAGCCCTCTCTGCCCAT-----CGAGCTGCACCCCG	4334
QY	1370	ACAAGTGGACCGTGACGCCCATCGAGCTGCCCGAAGGAGAGCTTGACCGTGAACGACA	1429
DB	4335	ACAAGTGGACCGTGACGCCCATCGTGCTGCCCGAAGGAGAGCTTGACCGTGAACGACA	4394
QY	1430	TCCAGAGCTGTGGGCAAGCTGAACGTGGGCGAGCCAGATCTACCCGGCATCAAGGTGC	1489
DB	4395	TCCAGAGCTGTGGGCAAGCTGAACGTGGGCGAGCCAGATCTACCCGGCATCAAGGTGC	4454
QY	1490	GCACGCTGTCAAGCTGTGCGCGGCGCCAGGCGCCTGACCGAGATCTGTGCCCTGACCG	1549
DB	4455	GCCAGCTGTCAAGCTGTGCGCGGCGCCAGGCGCCTGACCGAGATCTGTGCCCTGACCG	4514
QY	1550	AGGAGCGGAGCTTGGAGCTTGGCGGAGAACCGCGAGATCCTGCGGAGCCGCTGCACGGCG	1609
DB	4515	AGGAGCGGAGCTTGGAGCTTGGCGGAGAACCGCGAGATCCTGCGGAGCCGCTGCACGGCG	4574
QY	1610	TGTACTACGACCCAGGAGACTGTGTGGCGAGATCCAGAGAGCGGGCCACGACCACT	1669
DB	4575	TGTACTACGACCCAGGAGACTGTGTGGCGAGATCCAGAGAGCGGGCCACGACCACT	4634
QY	1670	GGACCTACCAAGATCTACAGGAGCCCTTCAGAACCTTGAACACCGCAAGTACGCCAAGA	1729
DB	4635	GGACCTACCAAGATCTACAGGAGCCCTTCAGAACCTTGAACACCGCAAGTACGCCCGCA	4694
QY	1730	TGCGCACCGCCACACACGACGTGAAGACAGCTGACCGAGGCGGTGCAGAGATCGCCA	1789
DB	4695	TGAAGGCGGCCACACACGACGTGAAGACAGCTGACCGAGGCGGTGCAGAGATCGCCA	4754
QY	1790	TGGAGAGCATCTGTGATCTGGGCGAAGACCCCAAGTTCCGCTTGCCTCCAGAGGAGA	1849
DB	4755	CCGAGAGCATCTGTGATCTGGGCGAAGACCCCAAGTTCCGCTTGCCTCCAGAGGAGA	4814
QY	1850	CTTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTGC	1909
DB	4815	CTTGGGAGGCTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTGC	4874
QY	1910	TGAACACCCCCCTTGTGTGAGCTGTGTACAGCTTGAAGAGGAGCCCATCATCGGC	1969
DB	4875	TGAACACCCCCCTTGTGTGAGCTGTGTACAGCTTGAAGAGGAGCCCATCATCGGC	4934
QY	1970	CCGAGACCTTCTACGTGGACGGCGCGCAACCCGCGAGACCAAGATCGGCAAGCGCGCT	2029

Db	4935	CGAGACCTTCTACGTGGACGGCGCCGCAACCGCGAGACCAAGCTGGGCAAGCGCGCT	4994
QY	2030	AGTGAGACCCAGCGGGCGCCGAGAAATCGTAGCCTGACCCGAGACACCAACGAGA	2089
Db	4995	ACGTGACCCAGCGCGCGCCGCAAGAGTGGTGCCCTGACCGACACCAACGAGA	5054
QY	2090	CGAGCTGCAGGCCATCCAGCTGGCCCTGCAGACAGCGGACGAGGTGAACATCGTGA	2149
Db	5055	CCGAGCTGCAGGCCATCCACCTGGCCCTGCAGACAGCGGCTGGAGGTGAACATCGTGA	5114
QY	2150	CCGACAGCCAGTACGCCCTGGGCGATCATCCAGGCCCGAGCCGACAGCGAGCGAGC	2209
Db	5115	CCGACAGCCAGTACGCCCTGGGCGATCATCCAGGCCCGAGCCGACAGCGAGCGAGC	5174
QY	2210	TGGTGAACCCAGATCATCGAGCAGCTGATCAAGNAGAGAGAGTGTACCTGAGCTGGTGC	2269
Db	5175	TGGTGAACCCAGATCATCGAGCAGCTGATCAAGNAGAGAGAGTGTACCTGAGCTGGTGC	5234
QY	2270	CGCCACCAAGGCGATCGCGCGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGCGATCC	2329
Db	5235	CGCCACCAAGGCGATCGCGCGCAACGAGCAGTGGTGAGCGCGCGCGATCC	5294
QY	2330	GCAAGGTGCTGTCTCGACGCGATCGATGGGGCGATCGTGATCTACCACTA	2381
Db	5295	GCAAGGTGCTGTCTCGACGCGATCGACAAGGCCCGAGGAGGACGAGAA	5346
RESULT	9		
ABK91617			
ID	ABK91617	standard; DNA; 9169 BP.	
XX			
AC	ABK91617;		
XX			
DT	14-AUG-2002	(first entry)	
XX			
XX		Modified HIV protein-encoding plasmid DNA #169.	
KW	HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV		
KW	Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;		
KW	acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.		
OS	Synthetic.		
XX			
PN	WO200232943-A2.		
XX			
PD	25-APR-2002.		
XX			
PF	14-AUG-2001; 2001WO-US25721.		
XX			
PR	14-AUG-2000; 2000US-225097P.		
PR	14-NOV-2000; 2000US-252115P.		
PR	28-MAR-2001; 2001US-279257P.		
XX			
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.		
PA	(CHAD/) CHADABARTI B K.		
XX			
PI	Nabel GJ, Huang Y;		
XX			
DR	WPI; 2002-452382/48.		
XX			
PT	New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and		
PT	its encoded protein, useful as vaccines for genetic or protein		
PT	immunisation for acquired immunodeficiency syndrome or human		
PT	immunodeficiency virus infection		
XX			
PS	Disclosure; Page 769-772; 794pp; English.		
XX			
CC	The invention relates to a nucleic acid molecule encoding a modified HIV		
CC	(human immunodeficiency virus) protein. The nucleic acid molecule and i		
CC	encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for		
CC	genetic or protein immunisation to a host, respectively. In particular		
CC	these are useful for ameliorating the disease.		

Db 4943 CCGAGACCTTCTACGTGGACGGCGCCCAACCGGAGACCAAGCTGGCAAGSCCGCT 5002
QY 2030 ACGTGACGACCGGGCCGCGAGAGATCGTGGAGCTGACCGAGACCAACACAGAGA 2089
Db 5003 ACGTGACGACCGGGCCGCGAGAGTGTGGCCCTTGACCGACACCAACACAGAGA 5062
QY 2090 CCGAGCTGACGAGCATCCAGCTGGCCCTGCAGGACACGCGAGGTGAACATCGTA 2149
Db 5063 CCGAGCTGACGAGCATCCACTGGCCCTGCAGGACACGCGCTGGAGGTGAACATCGTA 5122
QY 2150 CCGAGACCCAGTACGCCCTGGGATCATCCAGGCCACCGCCGACAGAGGAGAGCGAGC 2209
Db 5123 CCGAGACCCAGTACGCCCTGGGATCATCCAGGCCACCGCCGACAGAGGAGAGCGAGC 5182
QY 2210 TGGTGAACAGATCATCGACGAGTGTATCAAGAAGGAGAGGTGTACTGAGCTGGGTGC 2269
Db 5183 TGGTGAACAGATCATCGACGAGTGTATCAAGAAGGAGAGGTGTACTGAGCTGGGTGC 5242
QY 2270 CCGCCCAACAGGCGATCGCGCGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGCGATCC 2329
Db 5243 CCGCCCAACAGGCGATCGCGCGCAACGAGCAGGTGGACGGCTGGTGAGCGCCGGCATCC 5302
QY 2330 GCAAGGTGCTGTTCTCTGGAGCGCATCGATGGCGGCATCGTGATCTACAGTA 2381
Db 5303 GCAAGGTGCTGTTCTCTGGAGCGCATCGACAAAGGCCACGAGGAGGACAGAGAA 5354

RESULT 10
ABK91612
ID ABK91612 standard; DNA: 9194 BP.
XX
AC ABK91612;
XX
DT 14-AUG-2002 (first entry)
XX
DE Modified HIV protein-encoding plasmid DNA #164.
XX
KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX
OS Synthetic.
XX
PN W0200232943-A2.
XX
PD 25-APR-2002.
XX
PF 14-AUG-2001; 2001W0-US25721.
XX
PR 14-AUG-2000; 2000US-225097P.
PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADABARTI B K.
XX
PI Nabel GJ, Huang Y;
XX
XX WPI; 2002-452382/48.
XX
PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX
XX Disclosure; Page 753-756; 794pp; English.
XX
XX The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ Sequence 9194 BP; 2232 A; 2672 C; 1767 T; 0 other;
Query Match 80.4%; Score 1975; DB 24; Length 9194;
Best Local Similarity 90.9%; Pred. No. 1.2e-237;
Matches 2177; Conservative 0; Mismatches 190; Indels 28; Gaps 6;
QY 14 TGGCCGAGGCGCATGAGCCAGG-CCACCAGCGCCACATCCTGATGAGCGCAGCAACT 70
Db 2967 TGGCCGAGGCGCATGAGCCAGGTCACCAACAGCGCCACCATCATGATGAGCGCGGCAACT 3026
QY 71 TCRAAGGGCCCAACGCGCATCATCAAGTCTTCAACTGCGCAAGGAGGCGCACATCGGCC 130
Db 3027 TCSCCAACGAGCGCAAGATCGTGAAGTCTTCAACTGCGCAAGGAGGCGCACACCGCCC 3086
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCGCACACAGA 190
Db 3087 GCAACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCGCACACAGA 3146
QY 191 TGAAGGACTGCACCGAGCGCGCCAGGCCCAACTTCTTCCGCGAGGACCTGCCTTCCCCCAGG 250
Db 3147 TGAAGGACTGCACCGAGCGCGACAGGCTAA-TTTTATGAGGAAGATCTGGCCTTCCCACAG 3205
QY 251 GCAAGGGCCCGAGTTCGCCAGCGAGCAGAGAACCGCGCCCAACAGGCCACACAGCGCGCAGC 310
Db 3206 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGCCACAGCCCGCCACAGAGAGAGAGC 3265
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCCAGCGAGCGCGCGCGCGCGCGCGCGCA 364
Db 3266 TTCAGGTTTGGGGAAGAGACAACAACCTCCCTCTCAGAAGCAGGAGCGCATAGACAAGGAA 3325
QY 365 -----CCCTGAACCTTCCCGCAGATCACCTGTGGCAGCGCGCCCTGTGAGCATCANGG 418
Db 3326 CTGTATCTTTAGTCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAACAATAAGA 3385
QY 419 TGGCGCGCAGATCAAGGAGGCGCTGCTGGACACGCGCGCGCGCGCGCGCGCGCGCGCGCG 478
Db 3386 TAGGGGCCAGCTGAAGGAGCGCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3445
QY 479 AGATGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGATGCGCGGCGATCGCGCGCTTCATCA 538
Db 3446 AGATGAACCTGCCCGCGCGCTTGGAAAGCCCAAGATGATGCGCGGCGATCGCGCGCTTCATCA 3505
QY 539 AGGTGCGCCAGTACGACGAGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
Db 3506 AGGTGGGCGAGTACGACGAGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 3565
QY 599 TGTGATGCGCGCCCGCCCGCTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT 658
Db 3566 TGTGTTGGCGCCCGCCCGCTGACATCATCGCGCGCAACCTGCTGACCCAGATCGGCT 3625
QY 659 GCACCTGAACTTCCCGCATGACCCCGCTGAGACCGCTGCCCGTGAAGCTGAAGCCCGGCA 718
Db 3626 GCACCTGAACTTCCCGCATGACCCCGCTGAGACCGCTGCCCGTGAAGCTGAAGCCCGGCA 3685
QY 719 TGGACGCGCCCGCTGAGGAGGAGTGGCGCCCTGACCGAGGAGAGATCAGGCCCTTGACC 778
Db 3686 TGGACGCGCCCGCTGAGGAGGAGTGGCGCCCTGACCGAGGAGAGATCAGGCCCTTGACC 3745
QY 779 CCATCTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCCGAGAACCCCT 838
Db 3746 AGATCTGACGAGATGGAGAAGGAGGCAAGATCAGCAAGATCGGCCCGCCGAGAACCCCT 3805
QY 839 ACAACACCCCGCTTGGCCATCAAGAGAGAGGAGCAGCACCAAGTGGCGCGCAAGCTGGTGG 898
Db 3806 ACAACACCCCGCTTGGCCATCAAGAGAGGAGCAGCACCAAGTGGCGCGCAAGCTGGTGG 3865

QY 899 ACTTCCGCGAGCTGAACAACGCGACCCAGAGACTTCTGGGAGGTGCAGCTGGGCATCCCC 958
Db 3866 ACTTCCGCGAGCTGAACAACGCGACCCAGAGACTTCTGGGAGGTGCAGCTGGGCATCCCC 3925
QY 959 ACCCCGCGCCCTGAAGAAGAAAGAGCTGACCGTGTCTGGAGCTGGGGCGAGCCCTACT 1018
Db 3926 ACCCCGCGCCCTGAAGAAGAAAGAGCTGACCGTGTCTGGAGCTGGGGCGAGCCCTACT 3985
QY 1019 TCAGCTGCCCTTGGAGGAGGACTTCGCGAAGTACACCGCTTACCATCCCGAGCATCA 1078
Db 3986 TCAGCTGCCCTTGGAGGAGGACTTCGCGAAGTACACCGCTTACCATCCCGAGCATCA 4045
QY 1079 ACAACGAGACCCCGGATCCGCTACCAAGTACAACTGTCTGCCCGAGGGCTGGAAGGCA 1138
Db 4046 ACAACGAGACCCCGGATCCGCTACCAAGTACAACTGTCTGCCCGAGGGCTGGAAGGCA 4105
QY 1139 GCCCCAGCATCTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCCGGCCCGCAACC 1198
Db 4106 GCCCCAGCATCTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCCGGCCCGCAACC 4165
QY 1199 CCAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGCGAGCGACCTGGAGATCG 1252
Db 4166 CCAGATCGTGATCTACCAAGTACATGACCAACCTGTACGTGGCGAGCGACCTGGAGATCG 4225
QY 1253 GCCAGACCCCGGCAAGATCGAGAGCTGCGCAAGCAACCTGTCTGCGGTGGGGCTTCACCA 1312
Db 4226 GCCAGACCCCGGCAAGATCGAGAGCTGCGCAAGCAACCTGTCTGCGGTGGGGCTTCACCA 4285
QY 1313 CCCCCAGCAAGACGACCAAGAGGACCCCTTCTGCTGCCCAT-----CGAGCTGCACC 1366
Db 4286 CCCCCAGCAAGACGACCAAGAGGACCCCTTCTGCTGCCCATGCCGTACGAGCTGCACC 4345
QY 1367 CCAGCAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAAGAGAGAGCTGGACCGTGAACG 1426
Db 4346 CCAGCAAGTGGACCGTGCAGCCCATCGTGTCTGCCCGAAGAGAGAGCTGGACCGTGAACG 4405
QY 1427 ACATCAGAAAGCTGTGGGCAAGCTGAATGTGGGCGCAGCCAGATCTACCCCGGATCAAGG 1486
Db 4406 ACATCAGAAAGCTGTGGGCAAGCTGAATGTGGGCGCAGCCAGATCTACCCCGGATCAAGG 4465
QY 1487 TGGCCAGCTGTCAAGCTGTCTCGCGGGCGCCAAAGCCCTGACCGGACATCTGTCCTCCCTGA 1546
Db 4466 TGGCCAGCTGTCTCAAGCTGTCTCGCGGGCGCCAAAGCCCTGACCGGAGGTGGTCCCTGA 4525
QY 1547 CCAGAGGCGCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTCGCGAGCCCGTGCACG 1606
Db 4526 CCAGAGGCGCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTCGAGAGCCCGTGCACG 4585
QY 1607 GCCTGTACTAGACCCCGAGAGAGCTGTGTGGCCGAGATCTCAGAGAGCGAGCCAGAC 1666
Db 4586 GCCTGTACTAGACCCCGAGAGAGCTGTGTGGCCGAGATCTCAGAGAGCGAGCCAGAC 4645
QY 1667 AGTGGACCTACCAAGTCTACCAAGAGCCCTTCAAGAACCTGAGACCGGCAAGTACGCCA 1726
Db 4646 AGTGGACCTACCAAGTCTACCAAGAGCCCTTCAAGAACCTGAGACCGGCAAGTACGCCA 4705
QY 1727 AGATGGCGACCGCCACACCAACGAGCTGAAGCAGCTGACCGAGGCGGTGCAGAAATCG 1786
Db 4706 GCATGAAGGCGCCACACCAACGAGCTGAAGCAGCTGACCGAGGCGGTGCAGAAATCG 4765
QY 1787 CCATGGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCCGCTTCCGCTGCCCATCCAGAAG 1846
Db 4766 CCACCGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCCGCTTCCGCTGCCCATCCAGAAG 4825
QY 1847 AGACCTGGGAGACCTGTGTGACCGACTACTGGCAGCGCCACCTGGATCCCGAGTGGAGT 1906
Db 4826 AGACCTGGGAGACCTGTGTGACCGACTACTGGCAGCGCCACCTGGATCCCGAGTGGAGT 4885
QY 1907 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACCAAGTGGAGAGGAGGCCCATCATCG 1966
Db 4886 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACCAAGTGGAGAGGAGGCCCATCATCG 4945
QY 1967 GCGCCGAGACCTTCTAGTGGAGGCGCGCGCCCAACCCGCGAGACCAAGATCGGCAAGCGCG 2026

Db 4946 GCGCCGAGACCTTCTAGTGGAGCGCGCCCAACCGCGAGACCAAGCTGGCAAGGCCG 5005
QY 2027 GCTACGCTGACCGCGCGCGCGCGAGAGATCTGTAGCCCTGACCGAGACCAACCAACG 2086
Db 5006 GCTACGCTGACCGCGCGCGCGCGAGAGATCTGTAGCCCTGACCGAGACCAACCAACG 5065
QY 2087 AGACCCGAGCTGCGAGGCGCATCCAGCTGGCCCTGCGAGGACGCGGAGGTGAACATCG 2146
Db 5066 AGACCCGAGCTGCGAGGCGCATCCAGCTGGCCCTGCGAGGACGCGGAGGTGAACATCG 5125
QY 2147 TGACCGAGACCCAGTACGCCCTGGGCATCATTCAGGCCCGAGCCCGCAAGAGCGAGAGCG 2206
Db 5126 TGACCGAGACCCAGTACGCCCTGGGCATCATTCAGGCCCGAGCCCGCAAGAGCGAGAGCG 5185
QY 2207 AGCTGTGTAACCAAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACTGAGCTGGG 2266
Db 5186 AGCTGTGTAACCAAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACTGAGCTGGG 5245
QY 2267 TGCCCGCCCAAGAGGCGCATCGCGCGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGCGCA 2326
Db 5246 TGCCCGCCCAAGAGGCGCATCGCGCGCAACGAGCAGATCGACAAAGCTGGTGAGCGCCGCGCA 5305
QY 2327 TCAGCAAGGTGCTGCTTCTGGAGCGCATTCAGTGGCGGATCGTGATCTACCAAGTA 2381
Db 5306 TCAGCAAGGTGCTGCTTCTGGAGCGCATTCAGTGGCGGATCGACAAGGCCGAGGAGGAGAGAA 5360

RESULT 11
ABK91611
ID ABK91611 standard; DNA; 9194 BP.
XX
AC ABK91611;
XX
DT 14-AUG-2002 (first entry)
XX
DE Modified HIV protein-encoding plasmid DNA #163.
XX
KW HIV; human immunodeficiency virus; gene: ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX
OS Synthetic.
XX
PN WO200232943-A2.
XX
PD 25-APR-2002.
XX
PF 14-AUG-2001; 2001WO-US25721.
XX
PR 14-AUG-2000; 2000US-225097P.
PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADABARTI B K.
PI Nabel GJ, Huang Y;
XX
XX WPI; 2002-452382/48.
XX
PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX
PS Disclosure; Page 750-753; 794pp; English.
XX
CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ

Sequence 9194 BP; 2232 A; 2671 C; 2524 G; 1767 T; 0 other;

Query Match 80.3%; Score 1973.4; DB 24; Length 9194;
Best Local Similarity 90.9%; Pred. No. 1.9e-237;
Matches 2176; Conservative 0; Mismatches 191; Indels 28; Gaps 6;

QY 14 TGGCCGAGGCCATGAGCAGG---CCACCAGCGCCAAATCTCTGATGCGAGCAGCAACT 70
DB 2967 TGGCCGAGGCCATGAGCAGGTCGACCAAGCGCCACCATCATGATGACCGCGCAACT 3026

QY 71 TCAAGGGCCCAAGCGCATCATCAAGTGTTCACATGCGGCAAGAGGGCCACATCGCCC 130
DB 3027 TCCGCAACCAAGCGCATGCTGAAGTGTTCACATGCGGCAAGAGGGCCACACCGCCC 3086

QY 131 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTCGCGCAAGAGGGGCCACCA 190
DB 3087 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTCGCGCAAGAGGGGCCACCA 3146

QY 191 TGAAGGACTGCAACGAGCGCAGGCCAACTTCTTCGCGAGGACCTGGCCCTCCGCCAGG 250
DB 3147 TGAAGGACTGCAACGAGCGCAGGCCAACTTCTTCGCGAGGACCTGGCCCTCCGCCAGG 3205

QY 251 GCAAGGCCCGCGAGTTCCCGCAGGAGCAGCAACCGCGCAACAGCCCGCCAGCGCGAGC 310
DB 3206 GGAAGGCCCGCGAGTTTCTTCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3265

QY 311 TCGAGGTGCGGG-----CGACAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCGCA 364
DB 3266 TTCAGGTTTGGGGAAGAGACAACACTCCCTCTCAAGCAGGAGCGCGATAGACAAGGA 3325

QY 365 -----CCCTGAATCCCGCAGATACCTGCTGGCAGCGCCCTCTGGTGAGATCAAGG 418
DB 3326 CTGTATCTTTAGTTCCTTCAGATCACTCTTGGCAGCAGCCCTCTGTCACAAATAAGA 3385

QY 419 TGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGCG 478
DB 3386 TAGGGGCCAGCTGAAGGAGGCCCTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCG 3445

QY 479 AGATGACCTTGGCGGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
DB 3446 AGATGAACCTTGGCGGCGCGCTGGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 3505

QY 539 AGTGGCGCCAGTACGACAGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
DB 3506 AGTGGCGCCAGTACGACAGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 3565

QY 599 TGTGTATCGGCCCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCGAGCTGGGT 658
DB 3566 TGTGTGTGGCCCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCGAGCTGGGT 3625

QY 659 GCACCTTGAATTCGCCATCAGCCCATCGAGACCGTGCCTGGAAGTGAAGCCCGGCA 718
DB 3626 GCACCTTGAATTCGCCATCAGCCCATCGAGACCGTGCCTGGAAGTGAAGCCCGGCA 3685

QY 719 TGGAGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
DB 3686 TGGAGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGAGT 3745

QY 779 CCATCTCGAGGAGATGGAAGAGGCGCAAGATCAACAGATCGCGCGCGCGCGCGCGCG 838
DB 3746 AGATCTGCACCGAGATGGAAGAGGCGCAAGATCAACAGATCGCGCGCGCGCGCGCG 3805

QY 839 ACAACACCCCGGTTCGCCATCAAGAAAGAGACAGCACAAGTGGCGCAAGCTGGTG 898
DB 3806 ACAACACCCCGGTTCGCCATCAAGAAAGAGACAGCACAAGTGGCGCAAGCTGGTG 3865

QY 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTGAGCTGGGCATCCCCC 958
DB 3866 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTGAGCTGGGCATCCCCC 3925

QY 959 ACCCGCGCGCTGAAGAAGAAGAGCGTGAACGCTGCTGGACGTGGCGACGCGCTACT 1018
DB 3926 ACCCGCGCGCTGAAGAAGAAGAGCGTGAACGCTGCTGGACGTGGCGACGCGCTACT 3985

QY 1019 TCAGCGTCCCTCGACGAGGACTTCCGCAAGTACACGCCCTTACCATCCCGACATCA 1078
DB 3986 TCAGCGTCCCTCGACGAGGACTTCCGCAAGTACACGCCCTTACCATCCCGACATCA 4045

QY 1079 ACAACGAGACCCCGCGCATCCCTACCACTACAAGCTGCTGCCCGAGGCTGGAAGGCA 1138
DB 4046 ACAACGAGACCCCGCGCATCCCTACCACTACAAGCTGCTGCCCGAGGCTGGAAGGCA 4105

QY 1139 GCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCCGCCCGCAACC 1198
DB 4106 GCCCAGCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCCGAAGCAAGACC 4165

QY 1199 CCGAGATCGTGTATCTACCA-----GGCCCGCTGTAGTGGCGAGCACCTGGAGATCG 1252
DB 4166 CCGAGATCGTGTATCTACCACTGTAGTGCAGCATCTGTAGTGGCGAGCACCTGGAGATCG 4225

QY 1253 GCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGCTTCACCA 1312
DB 4226 GCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGCTTCACCA 4285

QY 1313 CCCCAGCAAGAAGCACCAGAAGGAGCCCGCTTCTGCGCCAT-----CGAGCTGCAACC 1366
DB 4286 CCCCAGCAAGAAGCACCAGAAGGAGCCCGCTTCTGTTGGATGGCTACGAGCTGCAACC 4345

QY 1367 CCGACAAGTGGACGCTGACGCCCATCGAGCTGCCGAGAGAGAGAGCTGGACCGTGAACG 1426
DB 4346 CCGACAAGTGGACGCTGACGCCCATCGAGCTGCCGAGAGAGAGAGCTGGACCGTGAACG 4405

QY 1427 ACATCCAGAAGCTGCTGGCAAGCTGAAGTGGCGCAGCAGATCTACCCCGCATCAAGG 1486
DB 4406 ACATCCAGAAGCTGCTGGCAAGCTGAAGTGGCGCAGCAGATCTACCCCGCATCAAGG 4465

QY 1487 TCGCCAGCTGTGCAAGCTGCTGCGCGCGCAAGGCCCTGACCGACATCTGTCGCCCTGA 1546
DB 4466 TCGCCAGCTGTGCAAGCTGCTGCGCGCGCAAGGCCCTGACCGAGGTGGTGGCCCTGA 4525

QY 1547 CCGAGGAGCCAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGACCGCTGACG 1606
DB 4526 CCGAGGAGCCAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGACCGCTGACG 4585

QY 1607 CGTGTACTTACACCCCGCAGCAGGACCTGCTGGCGGAGATCCAGAGCAGGCGCACGACC 1666
DB 4586 CGTGTACTTACACCCCGCAGCAGGACCTGATCGCGGAGATCCAGAGCAGGCGCACGACC 4645

QY 1667 AGTGAACCTTACAGATCTTACAGGAGCCCTTCAAGAAGCTGAAGACCGCGCAAGTACGCCA 1726
DB 4646 AGTGAACCTTACAGATCTTACAGGAGCCCTTCAAGAAGCTGAAGACCGCGCAAGTACGCCA 4705

QY 1727 AGATGGCACCCCGCCACACCAAGCAGCTGAAGCAGCTGACCGAGCGCGCTGAGAGATCG 1786
DB 4706 GCATGAAGGGCGCCCGCCACCAAGCAGCTGAAGCAGCTGACCGAGCGCGCTGAGAGATCG 4765

QY 1787 CCATGAGAGCATCTGATCTGGGCAAGACCCCGCAAGTTCGCGCTGCCCATCCAGAGG 1846
DB 4766 CCATGAGAGCATCTGATCTGGGCAAGACCCCGCAAGTTCGCGCTGCCCATCCAGAGG 4825

QY 1847 AGACCTTGGAGACCTTGGTGGACCGCTACTTGGCAGCGCACCTGGATCCCGAGTGGGAGT 1906
DB 4826 AGACCTTGGAGACCTTGGTGGACCGCTACTTGGCAGCGCACCTGGATCCCGAGTGGGAGT 4885

QY 1907 TCGTGAACACCCCGCTTGGTGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1966
DB 4886 TCGTGAACACCCCGCTTGGTGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4945

QY 1967 GCGCGGAGACCTTCTACGTGGAGCGCGCGCCCAACCGGAGACCAAGATCGCAAGGCGCG 2026

Db 4946 GCGCCGAGACCTTCTAGCTGGACGGCGCCGCAACCGCGAGACCAAGCTGGCAAGGCCG 5005
QY 2027 GCTACGTGACCGACCGCGCGCGCAGAGATCGTGGCCCTGACCGAGACCAACACAGA 2086
Db 5006 GCTACGTGACCGACCGCGCGCGCAGAGGTGTGCCCCCTGACCGACACCAACACAGA 5065
QY 2087 AGACCGAGCTGCAGGGCCATCCAGCTGGCCCTGCAGGACAGCGCGAGGTGAACATCG 2146
Db 5066 AGACCGAGCTGCAGGGCCATCCAGCTGGCCCTGCAGGACAGCGCGCTGGAGGTGAACATCG 5125
QY 2147 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCCAGCAAGAGCGAGAGCG 2206
Db 5126 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCCAGCAAGAGCGAGAGCG 5185
QY 2207 AGCTGGTGAACCCAGATCATCGACGACGTGATCAAGAGGAGAGGTGTACTGAGCTGGG 2266
Db 5186 AGCTGGTGAACCCAGATCATCGACGACGTGATCAAGAGGAGAGGTGTACTGAGCTGGG 5245
QY 2267 TGCCCGCCCAACAGGGGATCGCGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCA 2326
Db 5246 TGCCCGCCCAACAGGGGATCGCGCGGCAACGAGCAGGTGGAGCGCCTGGTGAGCGCGCA 5305
QY 2327 TCCGCAAGGTGCTTCTCGGACGGCATCGATGGCGGCATCGTGATCTACCACTA 2381
Db 5306 TCCGCAAGGTGCTTCTCGGACGGCATCGACAAGGCCAGGAGGACAGCA 5360

RESULT 12

ABK91613
ID ABK91613 standard; DNA; 12411 BP.

XX
AC ABK91613;

XX
DT 14-AUG-2002 (first entry)

XX
DE Modified HIV protein-encoding plasmid DNA #165.

XX
KW HIV; human immunodeficiency virus; gene; ds: circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX
OS Synthetic.

XX
PN W0200232943-A2.

XX
PD 25-APR-2002.

XX
PF 14-AUG-2001; 2001WO-US25721.

XX
PR 14-AUG-2000; 2000US-225097P.

PR
14-NOV-2000; 2000US-252115P.

PR
28-MAR-2001; 2001US-279257P.

XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (CHAD/) CHADRABARTI B K.

XX
PI Nabel GJ, Huang Y;

XX
WPI; 2002-452382/48.

XX
PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -

XX
PS Disclosure; Page 756-760; 794pp; English.

XX
CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

XX
SQ Sequence 12411 BP; 2938 A; 3733 C; 3476 G; 2364 T; 0 other;

Query Match 80.3%; Score 1973.4; DB 24; Length 12411;
Best Local Similarity 90.9%; Pred. No. 1.9e-237;
Matches 2176; Conservative 0; Mismatches 191; Indels 28; Gaps 6;

QY 14 TGCCCGAGGCCATGAGCGAGG---CCACCAGCGCCACATCTTGATCGACGCGCAACT 70
Db 2967 TGCCCGAGGCCATGAGCGAGGTGACCAACAGCGCCACCATCATGATCGACGCGGCACT 3026
QY 71 TCAAGGGCCCAAGCGGCATCATCAAGTGTCTCAACTCGCGCAAGGAGGGCCACATCGCCC 130
Db 3027 TCCGCAACCGACGCAAGATCGTGAAGTGTCTCAACTCGCGCAAGGAGGGCCACACGCC 3086
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGTGGAAGTGGCGAAGAGGGCCACCAGA 190
Db 3087 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGTGGAAGTGGCGAAGAGGGCCACCAGA 3146
QY 191 TGAAGGACTGCACCGAGCGCCGCAACTTCTTCCCGAGGACCTGGCCTTCCGCCAGG 250
Db 3147 TGAAGGACTGCACCGAGCGGACAGGCTTAA-TTTTTGAAGAGATCTGGCCTTCCCAAG 3205
QY 251 GCAAGGCCCGCGAGTTTCCCGAGCGAGCAACCGCGCCCAACAGCCCGCCACCGCGCAGC 310
Db 3206 GGAAGGCCAGGGAATTTTCTCAGAGCAGACAGAGCAACAGCCCGCCACCAAGAGAGC 3265
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGCGCGCAGGGCA 364
Db 3266 TTCAGGTTTGGGGAAGAGACAACATCTCCTCAGAAAGCAGGAGCGGTAGACAAGGAA 3325
QY 365 -----CCCTGAATTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCAGCCCTGCTCACAATAAGA 3385
QY 419 TGGCGCGCCAGATCAAGAGAGGCGCTTGTGACACCGCGCGCGCAGACACCGTGTGAGG 478
Db 3386 TAGGGGCGCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGCAGACACCGTGTGAGG 3445
QY 479 AGATGAGCCTGCCCGCGCAAGTGGAAAGCCCAAGATGATCGCGGCATCGCGGCTTCATCA 538
Db 3446 AGATGAACCTGCCCGCGCGCTGGAAAGCCCAAGATGATCGCGGCATCGCGGCTTCATCA 3505
QY 539 AGGTGCGCCAGTACGACCCAGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
Db 3506 AGGTGGCCAGTAGGACCCAGATCTGATCGAGATCTGCGGCCACAGGCCATCGGCACCG 3565
QY 599 TGTGATCGGCGCCCGCCCGTGAACATCATCGCGCGCCCAACATGCTGACCCAGCTGGCT 658
Db 3566 TGTGATCGGCGCCCGCCCGTGAACATCATCGCGCGCCCAACCTGCTGACCCAGATCGCT 3625
QY 659 GCACCTGAACTTCCCGCATCGACCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCA 718
Db 3626 GCACCTGAACTTCCCGCATCGACCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCA 3685
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 3686 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGCTGG 3745
QY 779 CCATCTGCGAGGAGATGGAGAAGGAGGCAAGATCAACAGATCGGCCCGCCGAGAACCCCT 838
Db 3746 AGATCTGCACCGAGATGGAGAAGGAGGCAAGATCAGCAAGATCGGCCCGCCGAGAACCCCT 3805
QY 839 ACAACACCCCGCTTTCGCCCATCAAGAGAGGACAGCACCACCAAGTGGCGCAAGCTGTGG 898
Db 3806 ACAACACCCCGCTTTCGCCCATCAAGAGAGGAGCAGCACCACCAAGTGGCGCAAGCTGTGG 3865

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

XX	Sequence 9785 BP; 2360 A; 2843 C; 2731 G; 1851 T; 0 other;	
QQ	Query Match 80.2%; Score 1970; DB 24; Length 9785;	
CC	Best Local Similarity 90.9%; Pred. No. 5.1e-237;	
XX	Matches 2174; Conservative 0; Mismatches 185; Indels 33; Gaps 6;	
QY	14 TGGCCGAGGCCATGAGCCAGGCCACCAGCCCAACATCTCTGATGCGAGCGCAACTTCA 73	
DB	2967 TGGCCGAGGCCATGAGCCAGGCTGAGCAGCACCAACATCATGATGCGAGCGCGCAACTTCA 3026	
QY	74 AGGGCCCCAAGCCATCATCAAGTCTTCAACTGCGGCAGAGGAGGCCACATGCCCGCA 133	
DB	3027 AGGGCCCCAAGCG---CATCAAGTCTTCAACTGCGGCAGAGGAGGCCACCTGGCCCGCA 3083	
QY	134 ACTGCCCGCCGCCCGCAAGAGGGTCTGGAAGTGCAGAGGAGGCCACAGATGA 193	
DB	3084 ACTGCCCGCCCGTGGCAAGAGGGCTGCTGGAAGTGCAGAGGAGGCCACAGATGA 3143	
QY	194 AGGACTGCAACGAGCGCCAGGCCAATCTTCTCCGAGGACCTTGGCCCTTCCCCAGGGCA 253	
DB	3144 AGGACTGCAACGAGCGACAGGCTAA-----TAGGGAAGATCTGGCCCTTCCCACAGGGA 3197	
QY	254 AGGCCCGCGAGTTCCCGAGCGAGCAGNACCGCGCCACACGCCCCACAGCGCGGAGTGC 313	
DB	3198 AGGCCAGGGAATTTCTTCTAGAGCAGACACAGAGCCCAACAGCCCCACCAAGAGAGGCTC 3257	
QY	314 AGGTGGCGCG- ----CGACAAACCCCGCAGCGAGCGCGCGCGAGCGCCAGGGCA-- 364	
DB	3258 AGGTTGGGAAGAGACACAACTCCCTCTCAGAGCAGAGGAGCGGATAGACAGGAAGTCTG 3317	
QY	365 ---CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTTGGTGAAGTCAAGSTGG 421	
DB	3318 TATCCTTTAGCTTCCCTCAGATCACCTTTTGGCAGCGACCCCTCGTCACATAAAGATAG 3377	
QY	422 GCGGCAGATCAAGGAGCCCTGCTGGACACCGCGCGCGCGAGCAGACCGTCTGGAGGAGA 481	
DB	3378 GGGCGAGCTGAAGAGGCGCCCTTCTAGACACCGCGCGCGAGCAGACCGTCTGGAGGAGA 3437	
QY	482 TGAGCTTGCCTGCGGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGGGCTTCAACAAG 541	
DB	3438 TGAACCTGCCCGCGCTGGAAGCCCAAGATGATCGCGGCATCGCGGGCTTCAACAAG 3497	
QY	542 TGGCGCAGTACGACAGATCCTGATCGAGATCTGCGGCAGAGAGGCCATCGGCACCGTGC 601	
DB	3498 TGGGCCAGTACGACAGATCCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGC 3557	
QY	602 TGATCGGCCACCCCGTGAACATCATCGCGCGCACATGCTGACCCAGCTGGGTGCA 661	
DB	3558 TGTGGGCCACCCCGTGAACATCATCGCGCGCACATGCTGACCCAGATCGCGTGA 3617	
QY	662 CCCTGAACTTCCCGCATCAGCGCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCATGG 721	
DB	3618 CCCTGAACTTCCCGCATCAGCGCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCATGG 3677	
QY	722 ACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGGCCCTGACCGGCCA 781	
DB	3678 ACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGGCCCTGGTGGAGA 3737	
QY	782 TCTGCCAGGAGATGGAGAGGAGGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTACA 841	
DB	3738 TCTGCCAGGAGATGGAGAGGAGGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTACA 3797	
QY	842 ACACCCCGGTGTGCGCATCAAGAAAGAGAGACAGCAGTGGCGCAGCTGGTGGACT 901	
DB	3798 ACACCCCGGTGTGCGCATCAAGAAAGAGAGACAGCAGTGGCGCAGCTGGTGGACT 3857	

QY	902 TCCGCGAGCTGAACAAAGCCACCCAGGACTTCTGGGAGGTGACGCTGGGATCCCCCACC 961	
DB	3858 TCCGCGAGCTGAACAAAGCCACCCAGGACTTCTGGGAGGTGACGCTGGGATCCCCCACC 3917	
QY	962 CCGCCGCTTGAAGAAAGAGCGTGCCTGCTGGAGGTGGGCGAGCCCTACTTCA 1021	
DB	3918 CCGCCGCTTGAAGAAAGAGCGTGCCTGCTGGAGGTGGGCGAGCCCTACTTCA 3977	
QY	1022 GCGTGGCCCTGGAGGAGCTTCCGCAAGTACACCGCTTCAACCATCCCGAGCATCAACA 1081	
DB	3978 GCGTGGCCCTGGAGGAGCTTCCGCAAGTACACCGCTTCAACCATCCCGAGCATCAACA 4037	
QY	1082 ACAGAGCCCCCGCATCCGCTACAGTACAAAGTGTGCCCCAGGGGTGGAAGGGCAGCC 1141	
DB	4038 ACAGAGCCCCCGCATCCGCTACAGTACAAAGTGTGCCCCAGGGGTGGAAGGGCAGCC 4097	
QY	1142 CCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCCGCCCGCAACCCCG 1201	
DB	4098 CCAGCATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCCCTTCCCGCAACCAACCCCG 4157	
QY	1202 AGATCGTGATCTACCA-----GGCCCCCTGTACGTGGGCGAGCGACTGGAGATCGGCC 1255	
DB	4158 ACATCGTGATCTACCAAGTACATGACCAACCTGTACGTGGGCGAGCGACTGGAGATCGGCC 4217	
QY	1256 AGACCCGCGCAAGATCGAGAGCTGCGCAAGCACTGTCTGCGCTGGGGTTCACACACC 1315	
DB	4218 AGACCCGCGCAAGATCGAGAGCTGCGCAAGCACTGTCTGCGCTGGGGTTCACACACC 4277	
QY	1316 CCACAGAGAGCACCAGAGGAGCCCTTCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1369	
DB	4278 CCACAGAGAGCACCAGAGGAGCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4337	
QY	1370 ACAAGTGCAGCGTGCAGCCCATCGAGCTGCGCGAGAGAGAGCTGACCGTGAACGACA 1429	
DB	4338 ACAAGTGCAGCGTGCAGCCCATCGTCTGCGCGAGAGAGAGCTGACCGTGAACGACA 4397	
QY	1430 TCCAGAGCTGGTGGGCAAGCTGAACTGGGCGAGCCAGATCTACCCCGCGCATCAAGTGC 1489	
DB	4398 TCCAGAGCTGGTGGGCAAGCTGAACTGGGCGAGCCAGATCTACCCCGCGCATCAAGTGC 4457	
QY	1490 GCCAGCTGTCAAGCTGCTGCGCGGCCCAAGGCCCTTGACCGAGATCTGCTCCCTGACCG 1549	
DB	4458 GCCAGCTGTCAAGCTGCTGCGCGGCCCAAGGCCCTTGACCGAGATCTGCTCCCTGACCG 4517	
QY	1550 AGAGGCCGAGCTGGAGCTGGCGAGAACCGCGAGATCTCTGCGCGAGCCGCTGCGACGGCG 1609	
DB	4518 AGAGGCCGAGCTGGAGCTGGCGAGAACCGCGAGATCTCTGCGCGAGCCGCTGCGACGGCG 4577	
QY	1610 TGTACTACGACCCAGCAAGGACTTGGTGGCGAGATCCAGAGAGAGGGCCACGACAGT 1669	
DB	4578 TGTACTACGACCCAGCAAGGACTTGTATCGCGAGATCCAGAGAGAGGGCCACGACAGT 4637	
QY	1670 GGACCTACAGATCTACCGAGGCCCTTCAAGACCTTGAAGACCGGCAAGTACGCCAAGA 1729	
DB	4638 GGACCTACAGATCTACCGAGGCCCTTCAAGACCTTGAAGACCGGCAAGTACGCCAAGA 4697	
QY	1730 TGGCGACCGCCACACCAAGCGTGAAGCAGCTGACCGAGCGCGTGCAGAAAGATCGCCA 1789	
DB	4698 TGAAGCGCGCCACACCAAGCGTGAAGCAGCTGACCGAGCGCGTGCAGAAAGATCGCCA 4757	
QY	1790 TGGAGAGCATCTGTATCTGGGGCAAGACCCCAAGTTTCCGCCCTGCCCATCCAGAAAGAGA 1849	
DB	4758 CCGAGAGCATCTGTATCTGGGGCAAGACCCCAAGTTTCAAGTTTCAAGTTTCAAGTTTCAAGTTT 4817	
QY	1850 CTTGGGAGACTTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCG 1909	
DB	4818 CTTGGGAGACTTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCG 4877	
QY	1910 TGAACACCCCGCTTGGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCTCGCG 1969	
DB	4878 TGAACACCCCGCTTGGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCTCGCG 4937	
QY	1970 CCGAGAGCTTCTACGTGGAGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCCCGCT 2029	

Db 4938 CCGAGACCTTCTAGCTGGACGGCCGCCCAACCGCGAGACCAAGCTGGGCAAGCCGGCT 4997
QY 2030 ACCTGACCGACCGCGCGGCGGAGAGATCGTGAGCCCTGACCGAGACCAACACAGAGA 2089
Db 4998 ACCTGACCGACCGCGGCGGAGAGATCGTGAGCCCTGACCGAGACCAACACAGAGA 5057
QY 2090 CCGAGCTGACAGGCATCCAGCTGGCCCTGTCAGGACAGCGGCGAGGTGAACATCGTGA 2149
Db 5058 CCGAGCTGACAGGCATCCAGCTGGCCCTGTCAGGACAGCGGCGAGGTGAACATCGTGA 5117
QY 2150 CCGACAGCCAGTACGCTGGGATCATCCAGGCGGCGGAGCGGCGGAGAGAGCGAGC 2209
Db 5118 CCGACAGCCAGTACGCTGGGATCATCCAGGCGGCGGAGCGGCGGAGAGAGCGAGC 5177
QY 2210 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACTGAGCTGGGTGC 2269
Db 5178 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACTGAGCTGGGTGC 5237
QY 2270 CCGCCCAAGAGGCGATCGGCGGCGGAGAGATCGACAAAGCTGGTGAAGGCGATCC 2329
Db 5238 CCGCCCAAGAGGCGATCGGCGGCGGAGAGATCGACAAAGCTGGTGAAGGCGATCC 5297
QY 2330 GCAAGGTGCTGCTTCTGGAGGCGATCGATGGCGCATCGTGATCTACAGTA 2381
Db 5298 GCAAGGTGCTGCTTCTGGAGGCGATCGACAAAGGCGGAGGAGGAGGAGGAGGAG 5349

RESULT 14

ABK91619
ID ABK91619 standard; DNA; 9167 BP.
XX AC ABK91619;
DT 14-AUG-2002 (first entry)
DE Modified HIV protein-encoding plasmid DNA #171.
XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX Synthetic.
OS
XX WO200232943-A2.
XX 25-APR-2002.
XX 14-AUG-2001; 2001WO-US25721.
XX 14-AUG-2000; 2000US-225097P.
PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX (USHS) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADRABARTI B K.
XX Nabel GJ, Huang Y;
XX WPI; 2002-452382/48.
XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX Disclosure; Page 775-778; 794pp; English.
XX

CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX

SQ Sequence 9167 BP; 2225 A; 2668 C; 2517 G; 1757 T; 0 other;

Query Match 80.08; Score 1966.6; DB 24; Length 9167;
Best Local Similarity 90.7%; Pred. No. 1.4e-236;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

QY 14 TGGCCGAGGCCATGAGCCAGG---CCACCAGCGCCCAACTCTGATGAGCGCAGCAACT 70

Db 2964 TGGCCGAGGCCATGAGCCAGGTCACACAGCGCCCAACTCTGATGAGCGCAGCAACT 3023

QY 71 TCAGAGGCCCCCAAGCGCATCATCAAGTCTTCAACTGCGGAAGAGGGCCACATCGGCC 130

Db 3024 TCCGCAACACGCGCAAGATCGTGAAGTCTTCAACTGCGGAAGAGGGCCACATCGGCC 3083

QY 131 GCAACTGCCGCGCCCGCCGCAAGAGAGGCTGCTGGAAGTGGCGGCAAGAGGGCCACAGA 190

Db 3084 GCAACTGCCGCGCCCGCCGCAAGAGAGGCTGCTGGAAGTGGCGGCAAGAGGGCCACAGA 3143

QY 191 TGAAGGACTGCACCGAGCGCCAGGCAACTTCTTCCGCGAGGACCTGGCTTCCCCAGG 250

Db 3144 TGAAGGACTGCACCGAGCGGACAGGCTAA-----TAGGGAAGATCTGGCTTCCACAAAG 3197

QY 251 GCAAGGCCCGGAGTTCCCGAGCGAGCAGAACCGCCCAACAGCCCGCCACCGCGCAGC 310

Db 3198 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGACCAACAGCCCGCCACCGAGAGAGC 3257

QY 311 TGCAGGTGCGGG-----CGACAACCCCGCAGGAGGCGCGCCCGAGCGCCAGGCGCA 364

Db 3258 TTCAGGTTGGGGAAGAGACAACAACTCCCTCTCAGAGCAGGAGCCGATAGACAAGGAA 3317

QY 365 -----CCCTGAATCTCCCGCAGATCACCTGTGCGAGCGCCCTGGTGAGCATCAAGG 418

Db 3318 CTGTATCTTTAGCTTCCCTCAGATCACTTTGGCAGCGACCCCTCTGCTCAATAAAGA 3377

QY 419 TGGCGCGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCAGACACCTGTGTGAGG 478

Db 3378 TAGGGGCGCAGCTGAAGGAGGCCCTTCTAGACACCGGCGCGCAGACACCTGTGTGAGG 3437

QY 479 AGATGAGCTCCCGCGCAAGTGGAGCCCAAGATGATCGGCGGCAATCGCGCGCTTCATCA 538

Db 3438 AGATGAACCTGCCCGCGCTGGAAGGCCCAAGATGATCGGCGGCAATCGCGCGCTTCATCA 3497

QY 539 AGGTGCGCGAGTACGACAGATCCTGATCGAGATCTGGGCAAGAAGGCCATCGGCACCG 598

Db 3498 AGGTGCGCGAGTACGACAGATCCTGATCGAGATCTGGGCGCAAGAAGGCCATCGGCACCG 3557

QY 599 TGCTGATCGGCGCCCGCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGTGGGCT 658

Db 3558 TGCTGATCGGCGCCCGCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGTGGGCT 3617

QY 659 GCACCTTGAACCTTCCCGATCAGCCCGATCGAGACCGTGGCGGCAATGAAGCCCGGCA 718

Db 3618 GCACCTTGAACCTTCCCGATCAGCCCGATCGAGACCGTGGCGGCAATGAAGCCCGGCA 3677

QY 719 TGGAGCGCCCAAGTGAAGCAGTGGCCCTTGACCGAGGAGAAGATCAAGGCCCTGACCG 778

Db 3678 TGGAGCGCCCAAGTGAAGCAGTGGCCCTTGACCGAGGAGAAGATCAAGGCCCTGACCG 3737

QY 779 CCATCTCGAGGAGATGAGAGAGGCGCAAGATCACCAAGATCGGCCCGCGAGAACCCCT 838

Db 3738 AGATCTGCACCGAGATGAGAGAGGCGCAAGATCACCAAGATCGGCCCGCGAGAACCCCT 3797

QY 839 ACAACACCCCGTGTTCGCCATCAAGAGAGGACAGACCAAGTGGCGCAAGCTGGTGG 898

Db 3798 ACAACACCCCGTGTTCGCCATCAAGAGAGGACAGACCAAGTGGCGCAAGCTGGTGG 3857

Qy	899	ACTTCGCGAGCTGAACAAGCCACCCAGGACTTCTGGAGGTCGACGTGGGCAATCCCC	958
Db	3858	ACTTCGCGAGCTGAACAAGCCACCCAGGACTTCTGGAGGTCGACGTGGGCAATCCCC	3917
Qy	959	ACCCGCCGCGCTGAAGAAGAAGACGCTGACCGTCTGCTGACGTCGGGAGCGCCTACT	1018
Db	3918	ACCCGCCGCGCTGAAGAAGAAGACGCTGACCGTCTGCTGACGTCGGGAGCGCCTACT	3977
Qy	1019	TCAGCGTGCCTTGGACGAGGACTTCCGCAAGTAGTACACGCCCTTCCACATCCCCAGCATCA	1078
Db	3978	TCAGCGTGCCTTGGACGAGGACTTCCGCAAGTAGTACACGCCCTTCCACATCCCCAGCATCA	4037
Qy	1079	ACAACGAGACCCCCGGCATCCCGTACCAAGTACACAGTCTGCTGCCACAGGCGTGAAGGGCA	1138
Db	4038	ACAACGAGACCCCCGGCATCCCGTACCAAGTACACAGTCTGCTGCCACAGGCGTGAAGGGCA	4097
Qy	1139	GCCCCAGCATTTCCAGACGACATGACCAAGATCCTGGAGCCCTTCCGCGCCGCGCAACC	1198
Db	4098	GCCCCAGCATTTCCAGTGCAGATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGAACCC	4157
Qy	1199	CCGAGATCGTGATCTACCA-----GGCCCCCTGTAGCTGGGAGGAGCACTGGAGATCG	1252
Db	4158	CCGACATCGTGATCTACCAAGTACATGGACCACTGTAGCTGGGAGGAGCACTGGAGATCG	4217
Qy	1253	GCCAGACCGCGCCCAAGATCGAGGAGCTGCGCAAGCACTGCTGCGTGGGGCTTCACCA	1312
Db	4218	GCCAGACCGCACCAAGATCGAGGAGCTGCGCAAGCACTGCTGCGTGGGGCTTCACCA	4277
Qy	1313	CCCCCGACAGACACCAAGAGGAGCCCTTCTTCCCTCAT-----CGAGCTGCACC	1366
Db	4278	CCCCCGACAGAGACCAAGAGGAGCCCTTCTTCTGTGATGGGTACGAGCTGCACC	4337
Qy	1367	CCGACAAGTGGACCGTGCAGCCATCGAGCTGCCCGAAGAGAGAGCTGGACCGTGAACG	1426
Db	4338	CCGACAGTGGACCGTGCAGCCATCGTCTGCCCGAAGAGAGAGCTGGACCGTGAACG	4397
Qy	1427	ACATCCAGAAGCTGTGTGGCAAGCTGAACCTGGGCGACGCCAGATCTACCCGGCATCAAGG	1486
Db	4398	ACATCCAGAAGCTGTGTGGCAAGCTGAACCTGGGCGACGCCAGATCTACCCGGCATCAAGG	4457
Qy	1487	TGCGCAGCTGTCAAGCTGTCCGCGGCCCAAGGCCCTTACCGACATCGTGCCCTGTA	1546
Db	4458	TGCGCAGCTGTGTCAAGCTGTCCGCGGCCCAAGGCCCTTACCGAGGTGTGCCCCCTGA	4517
Qy	1547	CCGAGGAGCGGAGCTGGAGCTGGCGAGAACCGCGAGATCTTCGGGAGCCCGTGCACG	1606
Db	4518	CCGAGGAGCGGAGCTGGAGCTGGCGAGAACCGCGAGATCTTCGAGGAGCCCGTGCACG	4577
Qy	1607	GCGTGTACTACGACCCCAAGAGACCTTGTGGCGGAGATCCAGAAGCAGGCCACGACC	1666
Db	4578	GCGTGTACTACGACCCCAAGAGACCTTGTGGCGGAGATCCAGAAGCAGGCCACGACC	4637
Qy	1667	AGTGGACCTACAGATCTACCGAGGACCTTCAAGAACTGAAACCGGCAAGTACGCCA	1726
Db	4638	AGTGGACCTACAGATCTACCGAGGACCTTCAAGAACTGAAACCGGCAAGTACGCCC	4697
Qy	1727	AGATGGCACCGCCCAACCAAGAGCTGAAGCACTTCACCGAGCGCTGCAGAAAGATCG	1786
Db	4698	GCATGAAGGGCGCCCAACCAAGAGCTGAAGCACTGACCGAGCCGTGCAGAAAGATCG	4757
Qy	1787	CCATGGAGAGCATCGTGATCTGGGGCAAGACCCCCCAAGTTCGCGCTGCCATCCAGAAAGG	1846
Db	4758	CCACCGAGAGCATCGTGATCTGGGGCAAGACCCCCCAAGTTCGAGCTGCCATCCAGAAAGG	4817
Qy	1847	AGACCTGGGAGACCTGGTGGACCGACTACTGCGAGGCCACCTGTGATCCCGAGTGGGAGT	1906
Db	4818	AGACCTGGGAGGCTGGTGGACCGAGTACTGTGCAGGCCACCTGTGATCCCGAGTGGGAGT	4877
Qy	1907	TCGTGACACCCCCCTTGGTGAAGCTGTGTACCACTGGAGAGGAGGCCCATCATCG	1966
Db	4878	TCGTGAAACACCCCCCTTGGTGAAGCTGTGTGTACCACTGGAGAGGAGGCCCATCATCG	4937
Qy	1967	GCGCCGAGACCTTCTACGTGGAGCGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCGC	2026

Db	4938	GCGCCGAGACCTTCTACGTGGACGGCGCCGCAACCGCGAGACCAAGCTGGGCAAGGCCG	4997
Qy	2027	GCTACGTGACCCAGCCGGCGCGGCAAGAGATCGTGAGCCTGACCGAGACCAACCAACGAGA	2086
Db	4998	GCTACGTGACCGACCGCGCGCGGCAAGAGGTGGTGCCCTTGACCGACCAACCAACGAGA	5057
Qy	2087	AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGAGGAGGTGAACATCG	2146
Db	5058	AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCTGGAGGTGAACATCG	5117
Qy	2147	TGACCGACAGCCAGTACGCCCTTGGGCATCATCCAGGCCACGCCGACAGAGCGAGAGCG	2206
Db	5118	TGACCGACAGCCAGTACGCCCTTGGGCATCATCCAGGCCACGCCGACAGAGCGAGAGCG	5177
Qy	2207	AGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGG	2266
Db	5178	AGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGG	5237
Qy	2267	TGCCCCGCCACAAAGGCGATCGCGGCGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCA	2326
Db	5238	TGCCCCGCCACAAAGGCGATCGCGGCGCAACGAGCAGTGGAGCGGCTGGTGAGCGCGCA	5297
Qy	2327	TCGCGAAGGTGCTGTTCTTGGACGCGCATCGATGGCGGCGATCGTGATCTACCAGTA	2381
Db	5298	TCGCGAAGGTGCTGTTCTTGGACGCGCATCGACAAGGCCACGAGGAGGACGACAGAA	5352
RESULT 15			
ABK91614			
ID	ABK91614 standard; DNA; 9170 BP.		
XX	AC		
XX	ABK91614;		
XX	DT		
DT	14-AUG-2002 (first entry)		
XX	Modified HIV protein-encoding plasmid DNA #166.		
DE	HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HI-		
KW	Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;		
KW	acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.		
KW	Synthetic.		
OS	WO200232943-A2.		
XX	25-APR-2002.		
PN	14-AUG-2001; 2001WO-US25721.		
PD	14-AUG-2000; 2000US-225097P.		
XX	14-NOV-2000; 2000US-252115P.		
XX	28-MAR-2001; 2001US-279257P.		
XX	(USSH) US DEPT HEALTH & HUMAN SERVICES.		
PA	(CHAD/) CHADABARTI B K.		
PA	Nabel GJ, Huang Y;		
XX	WPI; 2002-452382/48.		
PI	New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and		
DR	its encoded protein, useful as vaccines for genetic or protein		
PT	immunisation for acquired immunodeficiency syndrome or human		
PT	immunodeficiency virus infection -		
XX	Disclosure; Page 760-763; 794pp; English.		
XX	The invention relates to a nucleic acid molecule encoding a modified H		
CC	(human immunodeficiency virus) protein. The nucleic acid molecule and		
CC	encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for		
CC	genetic or protein immunisation to a host, respectively. In particular		
CC	these are useful for ameliorating the symptoms of acquired		

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

XX
SQ Sequence 9170 BP; 2225 A; 2669 C; 2519 G; 1757 T; 0 other;

Query Match 80.0%; Score 1966.6; DB 24; Length 9170;
Best Local Similarity 90.7%; Pred. No. 1.4e-236;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

QY 14 TGGCCGAGGCGATGAGCCAGG---CCACCAGCGCCACATCTCTGATCAGCGGCAACT 70
DB 2967 TGGCCGAGGCGATGAGCCAGGTCACCAACAGCGCCACCATCATGATCAGCGGCAACT 3026
QY 71 TCAAGGCGCCAGCGATCATCAAGTGTCTTCAACTCGGCAAGGAGGCGGCATCGGCC 130
DB 3027 TCCGCAACCAAGCGCAAGATCGTGAAGTGTCTCAACTCGGCAAGGAGGCGGCACACCGGCC 3086
QY 131 GCAACTCGCGCGCCCGCCCAAGAGGCGTCTGGAAGTGGCGCAAGGAGGCGGCACACAGA 190
DB 3087 GCAACTCGCGCGCCCGCCCAAGAGGCGTCTGGAAGTGGCGCAAGGAGGCGGCACACAGA 3146
QY 191 TGAAGGACTGCACCGAGCGCCAGCCCAACTTCTTCCCGGAGGACTTGGCCCTTCCCGCAGG 250
DB 3147 TGAAGGACTGCACCGAGCGACAGGCTAA-----TAGGGAAGATCTGGCCCTTCCCAACAG 3200
QY 251 GCAAGGCGCGGAGTTCCCGCAGCGAGCAGACACCGCGCCACACGCCCCACAGCGCGAGC 310
DB 3201 GGAAGGCGCAGGAGTTTCTTACAGAGCAGACACAGAGCAACAGCGCCACACAGAGAGAGC 3260
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGCGCGCAGGCGCA 364
DB 3261 TTCAGGTTTGGGGAAGAGACAACACTCCCTCTCAGAAGCAGGAGCGCATAGACAAGGAA 3320
QY 365 -----CCCTGAACTTCCCGCAGATCACTCTGTGGCAGCGCCCTGTGTGAGCATCAAGG 418
DB 3321 CTGTATCTTCTAGCTTCCCTCAGATCACTCTTGTGCGAGCGACCCCTCTCACATANAAGA 3380
QY 419 TGGCGCGCGAGATCAAGGAGCGCCTGTGTGACACCGCGCGCGAGCAGCACCGTGTGAGG 478
DB 3381 TAGGGGCGCGAGCTGAAGAGGCGCCTTCTAGACACCGCGCGCGAGCAGCACCGTGTGAGG 3440
QY 479 AGATGAGCCTTCCCGCGCAAGTGGGAAGCCCAAGATGATCGCGCGCATCGCGGCTTCATCA 538
DB 3441 AGATGAGCCTTCCCGCGCGCTGGGAAGCCCAAGATGATCGCGCGCATCGCGGCTTCATCA 3500
QY 539 AGGTGCGCCAGTAGCAGACAGATCTGTATCGAGATCTCGGCAAGAGCGCATCGGCAACCG 598
DB 3501 AGGTGCGCCAGTAGCAGACAGATCTGTATCGAGATCTCGGCAAGAGCGCATCGGCAACCG 3560
QY 599 TGTGATCGGCGCCACCGCGTGAACATATCGCGCGCGACATCTGTGACCGAGTGGGCT 658
DB 3561 TGTGTTGGGCGCCCGCCCGTGAACATATCGCGCGCGACATCTGTGACCGAGTGGGCT 3620
QY 659 GCACCTTGAATCTCCCATCAGCCCGCATCGAGACCGTGCCTGAAGCTGAAGCGCGGCA 718
DB 3621 GCACCTTGAATCTCCCATCAGCCCGCATCGAGACCGTGCCTGAAGCTGAAGCGCGGCA 3680
QY 719 TGAAGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGGAGGAGAGATCAAGGCGCTGACCG 778
DB 3681 TGAAGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGGAGGAGAGATCAAGGCGCTGTTGG 3740
QY 779 CCATCTCGGAGGAGATGGAAGAGGCGGCAAGATCAACCAAGATCGGCGCGGCAAGCCCT 838
DB 3741 AGATCTGCACCGAGATGGAGAAGGAGGCGCAAGATCAGCAAGATCGGCGCGGCAAGCCCT 3800
QY 839 ACACACCGCGTGTTCGCCATCAAGAGAGGAGACACCAAGTGGCGGCAAGTGTGG 898
DB 3801 ACACACCGCGTGTTCGCCATCAAGAGAGGAGACACCAAGTGGCGGCAAGTGTGG 3860

QY 899 ACTTCCCGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTGACGTGGGCATCCCCC 958
DB 3861 ACTTCCCGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTGACGTGGGCATCCCCC 3920
QY 959 ACCCGCGCGGCTGAAGAAGAGAGAGCGTGCCTGCTGAGCGTGGGCGACCGCTACT 1018
DB 3921 ACCCGCGCGGCTGAAGAAGAGAGAGCGTGCCTGCTGAGCGTGGGCGACCGCTACT 3980
QY 1019 TCAGCGTCCCGCTGAGCAGGAGCTTCCGCAAGTACACCGCCTTCCACATCCCGAGCATCA 1078
DB 3981 TCAGCGTCCCGCTGAGCAGGAGCTTCCGCAAGTACACCGCCTTCCACATCCCGAGCATCA 4040
QY 1079 ACAACGAGAGACCCCGGATCCGCTACCAAGTACAACTGCTGCTCCCGAGGCTGGAGGCGCA 1138
DB 4041 ACAACGAGAGACCCCGGATCCGCTACCAAGTACAACTGCTGCTCCCGAGGCTGGAGGCGCA 4100
QY 1139 GCCCGAGATCTTCCAGAGCAGCATGACCAAGATCTTGGAGCGCCTTCCGCGCCCGCAAC 1198
DB 4101 GCCCGAGATCTTCCAGTGCAGCATGACCAAGATCTTGGAGCGCCTTCCGCGCCCGCAAC 4160
QY 1199 CCGAGATCGTATCTACCA-----GGCCCGCTGTAGTGGGAGCGACCTGGAGATCG 1252
DB 4161 CCGAGATCGTATCTACCAAGTACATGGACCACTGTAGTGGGAGCGACCTGGAGATCG 4220
QY 1253 GCCAGACCGCGCCAAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGTGGGCGCTTCAACA 1312
DB 4221 GCCAGACCGCACCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGTGGGCGCTTCAACA 4280
QY 1313 CCGCGAGACAGAGCAGCAAGAGAGCGCCCTTCTTCTGATGGGCTTACGAGCTGCACC 1366
DB 4281 CCGCGAGACAGAGCAGCAAGAGAGCGCCCTTCTTCTGATGGGCTTACGAGCTGCACC 4340
QY 1367 CCGACAAGTGCAGCGTGCAGCGCATCGAGCGCCCGGAGAGAGAGCTGGACCGTGAACG 1426
DB 4341 CCGACAAGTGCAGCGTGCAGCGCATCGAGCGCCCGGAGAGAGAGCTGGACCGTGAACG 4400
QY 1427 ACATCCAGAACTGTTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGCATCAAG 1486
DB 4401 ACATCCAGAACTGTTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGCATCAAG 4460
QY 1487 TGGCGCAGCTGTCAAGCTGCTGGCGCGCGCAAGGCGCCTGACCGACATCTGTGCCCTGA 1546
DB 4461 TGGCGCAGCTGTCAAGCTGCTGGCGCGCGCAAGGCGCCTGACCGACATCTGTGCCCTGA 4520
QY 1547 CCGAGAGGCGCGAGCTGGAGCTGGCGCGCAGACCGGAGATCTCGCGAGCGCCGTCACG 1606
DB 4521 CCGAGAGGCGCGAGCTGGAGCTGGCGCGCAGACCGGAGATCTCGCGAGCGCCGTCACG 4580
QY 1607 GCGTGTACTAGCAGCCCGCAGCAAGGACCTGTGTGGCGGAGATCCAGAGCAGGCGCCACG 1666
DB 4581 GCGTGTACTAGCAGCCCGCAGCAAGGACCTGTGTGGCGGAGATCCAGAGCAGGCGCCACG 4640
QY 1667 AGTGGACCTACCAAGATCTACCAAGGAGCGCTTCAAGACCTGAAGACCGCGCAAGTACG 1726
DB 4641 AGTGGACCTACCAAGATCTACCAAGGAGCGCTTCAAGACCTGAAGACCGCGCAAGTACG 4700
QY 1727 AGATGCGCGCCCGCCACACCAAGCAGCTGAAGCAGCTGACCGAGCGCGTGCAGAGATCG 1786
DB 4701 GCATGAAGGCGCGCCCGCCACCAAGCAGCTGAAGCAGCTGACCGAGCGCGTGCAGAGATCG 4760
QY 1787 CCATGGAGAGCATCTGATCTGGGCGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAGG 1846
DB 4761 CCACCGAGAGCATCTGATCTGGGCGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAGG 4820
QY 1847 AGACCTGGGAGACCTGTGTGAGCGGAGCTACTGTGAGCGCCACCTGTGATCCCGAGTGGAGT 1906
DB 4821 AGACCTGGGAGCGCTGTGTGAGCGGAGCTACTGTGAGCGCCACCTGTGATCCCGAGTGGAGT 4880
QY 1907 TCGTGAACACCCCGCTGTGTGAGCGTGTGATCGGAGCGCGTGTGATCCCGAGCGCCCATCATCG 1966
DB 4881 TCGTGAACACCCCGCTGTGTGAGCGTGTGATCGGAGCGCGTGTGATCCCGAGCGCCCATCATCG 4940
QY 1967 GCGCGGAGACCTTCTAGTGGAGCGCGCGCGCCCAACCGCGAGACCAAGATCGCAAGGCGCG 2026

```
|||||
Db 4941 GCCTCCGAGACCTTTCTACGTGGAGCGGCCCAACCGCGAGACCAAGCTGGGCAAGGCCG 5000
QY 2027 GCTACGTGACCGACCGGGCCGCGAGAAGATCGTGAGCCCTGACCGAGACCAACCAACGAGA 2086
Db 5001 GCTACGTGACCGACCGGGCCGCGAGAAGTGGTGGCCCTGACCGAGACCAACCAACGAGA 5060
QY 2087 AGACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACACGCGAGGTGAACATCG 2146
Db 5061 AGACCGAGCTGCAGGCGCATCCACCTGGCCCTGCAGGACACGCGGCTGGAGGTGAACATCG 5120
QY 2147 TGACCGACAGCCAGTAGCGCCCTGGGCATCATCCAGGCCCAAGCCGACAAAGAGCGAGGCG 2206
Db 5121 TGACCGACAGCCAGTAGCGCCCTGGGCATCATCCAGGCCCAAGCCGACAAAGAGCGAGGCG 5180
QY 2207 AGCTGGTGAACCAAGATCATCGACGAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 2266
Db 5181 AGCTGGTGAACCAAGATCATCGACGAGCTGATCAAGAGGAGAGGTGTACCTGGCCCTGGG 5240
QY 2267 TGCCCGCCCAACAAGGGCATCGCGGCAACGACGAGATCGGACAAGCTGGTGAGCAAGGGCA 2326
Db 5241 TGCCCGCCCAACAAGGGCATCGCGGCAACGACGAGGTGGACGGCCCTGGTGAGCGCGCGCA 5300
QY 2327 TCCGCAAGGTGCTGTTCCTGGACGGCATCGATGGCGGCATCGGTGATCTACCACTA 2381
Db 5301 TCCGCAAGGTGCTGTTCCTGGACGGCATCGACAAGGCCCAAGGAGGAGGACACGAGAA 5355
```

Search completed: February 10, 2003, 12:45:42
Job time : 582.804 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 11:44:48 : Search time 3377.75 Seconds
(without alignments)
11780.710 Million cell updates/sec

Title: US-09-610-313-32

Perfect score: 2457

Sequence: 1 gtcgacgcaccatggcga.....ggctagcaccggtgaattc 2457

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90.2	3.7	2598	11 AY103647	AY103647 Zea mays
2	85.6	3.5	951	13 BM321451	BM321451 rockefell
3	80.2	3.3	1132	13 BM320864	BM320864 rockefell
4	79.6	3.2	1165	13 BM320900	BM320900 rockefell
5	79	3.2	867	13 BM321430	BM321430 rockefell
6	76.6	3.1	1550	13 BM321022	BM321022 rockefell

7	75	3.1	545	13	BI724851	BI724851
8	75	3.1	862	13	BM321023	BM321023 rockefell
9	73.2	3.0	853	13	BM321393	BM321393 rockefell
10	72.8	3.0	566	13	BM587428	BM587428 170006873
11	72.2	2.9	640	10	BE601575	BE601575 HVSMEH009
12	71.6	2.9	500	13	BM372120	BM372120 EBrO03_SQ
13	71.6	2.9	538	13	BM368580	BM368580 EBrO08_SQ
14	71.6	2.9	540	9	AJ471121	AJ471121 AJ471121
15	71.6	2.9	566	14	BQ464692	BQ464692 HF02P20r
16	71.6	2.9	578	13	BM377112	BM377112 EBrO05_SQ
17	71.6	2.9	579	12	BF253640	BF253640 HVSMEF000
18	71.6	2.9	582	13	BM372871	BM372871 EBrO04_SQ
19	71.6	2.9	588	14	BQ765498	BQ765498 EBrO03_SQ
20	71.6	2.9	600	9	AL508778	AL508778 AL508778
21	71.6	2.9	625	14	BQ768295	BQ768295 EBrO08_SQ
22	71.6	2.9	636	13	BM370222	BM370222 EBrO08_SQ
23	71.6	2.9	660	14	BQ762385	BQ762385 EBrO01_SQ
24	71.6	2.9	700	9	AL508513	AL508513 AL508513
25	71.6	2.9	789	12	BF626988	BF626988 HVSMEB000
26	71.6	2.9	789	12	BF626988	BF626988 HVSMEB000
27	71.6	2.9	822	12	BG417596	BG417596 HVSMEK001
28	71.4	2.9	730	12	BF260696	BF260696 HVSMEF002
29	71.2	2.9	646	13	BM645449	BM645449 170006873
30	71.2	2.9	724	13	BM584191	BM584191 170006872
31	71	2.9	925	17	CNS0091P	AL053013 Drosophil
32	70.6	2.9	861	10	BE636696	BE636696 rockefell
33	70	2.8	494	12	BG349122	BG349122 947029E04
34	70	2.8	571	12	BF266979	BF266979 HV_CEA001
35	70	2.8	609	13	BJ292939	BJ292939 BJ292939
36	70	2.8	613	13	BJ226790	BJ226790 BJ226790
37	70	2.8	616	13	BJ280623	BJ280623 BJ280623
38	70	2.8	619	13	BJ291714	BJ291714 BJ291714
39	70	2.8	627	12	BG300932	BG300932 HVSMEB001
40	70	2.8	635	13	BJ290451	BJ290451 BJ290451
41	70	2.8	643	13	BJ289971	BJ289971 BJ289971
42	70	2.8	652	13	BJ290578	BJ290578 BJ290578
43	70	2.8	659	13	BJ226590	BJ226590 BJ226590
44	70	2.8	660	13	BJ288689	BJ288689 BJ288689
45	70	2.8	662	13	BJ281217	BJ281217 BJ281217

ALIGNMENTS

RESULT 1
AY103647
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AY103647
Zea mays
AY103647
AY103647.1
HTC
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2598)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whittitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 2598)
Coe,E.C.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
Location/Qualifiers
1. 2598
/organism="Zea mays"
/db_xref="MaizeDB:638378"
/db_xref="taxon:4577"
/clone="PC0142084"

2598 bp mRNA linear HTC 25-MAY-2002

/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"									
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"									
BASE COUNT	490 a	907 c	795 g	406 t					
ORIGIN									
Query Match	3.7%	Score 90.2;	DB 11;	Length 2598;					
Best Local Similarity	42.2%;	Pred. No. 2e-05;							
Matches 814;	Conservative 0;	Mismatches 1098;	Indels 15;	Gaps					
QY	296	CCACGACGCGGAGCTGCAGTGGCGGCGACAAACCCCGCAGCGAGGCGCGCGCGAGC	355						
DB	51	CCACACCTCCACCTCTGCGACCGCGGCGGCCCAACCAACACACCGGAGCGACGAA	110						
QY	356	GCCAGGCGACCTGAACTTCCCGCCAGATCAACCTGTGGCAGCGCCCTGGTGAGCATCA	415						
DB	111	TGGCGGCAACGGCGCATCTGTGAGAGCGACCCGCTGAACTGGGCGCGCGCGCGCG	170						
QY	416	AGTGGGCGGCGAGATCAAGAGGCGCTGTGGACACCGGGCGCGAGACACCGTGTG	475						
DB	171	AGCTGGCGGAGCGCACCTGTGACGAGGTGAAGCGCATGTGTGGCGCAGCGCGCAGCCG	230						
QY	476	AGGAGATGAGCCTGCCCGGAAGTGAAGGCCCAAGATGATCGGGCGCATCGCGGCTTCA	535						
DB	231	TGGTCAAGATCGAGGTTCCACCTTCGCGTCGCCAGGTGGCGCGCGCTGCCTTCGCCCA	290						
QY	536	TCAAGTGGCGCGAGTACGACGAGATCCTGTATCGAGATCTGCGGCAAGAAGGCCATCGGCA	595						
DB	291	AGGACGCGTCCGCGTCCGCGTCGAGTCGACGAGGAGGCGCCCGCGCGTCAAGGCCA	350						
QY	596	CGTGTGATCGGCG	655						
DB	351	GCAGCGAGTGGATCCTCGACTGTGATCGCGCGCGCGCGCGCATCTACGGGCTCACCA	410						
QY	656	GCTGCAACCTTGAATTCCTCCCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	715						
DB	411	GCTTGGCGGCGACCTCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	470						
QY	716	GATGAGCGGCG	775						
DB	471	TCAGGATCTCAACCGCGGAATCTTCGGCACCGCGCGCGCGCGCGCGCGCGCGCGCG	530						
QY	776	CGCGCATCTGGAGGAGATGGAGAAGGAGGCGAGATCACCAGATCGGCGCGCGGAA	835						
DB	531	AGTCAACCG	590						
QY	836	CCTACAACACCCCGCTTTCGCCATCAAGAAGAGGAGCAGCAACCAAGTGGCGCAAGCTGG	895						
DB	591	TCCGCTTCGAGATCTCGAGGCGCATCAGAGCTGTCTCAACCCGGTGTGAGCCCTGCC	650						
QY	896	TGAGCTTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGAGCTGGGCGATCC	955						
DB	651	TGCGCGCTCGGGGCGACCATCACCGCTGCGGCGACCTGGTCCCGCTCTCTACATCGCG	710						
QY	956	CCACCGCGCGCGCTGAAGAAGAGAGCGCTGACCGTGTGAGGTGGGCGAGCGCT	1015						
DB	711	GCCTCATACGGGCG	767						
QY	1016	ACTTCAAGCTGCGCGTGGAGGAGCTTCGCGAAGTACACCGCTTCAACCATCCCGAGCA	1075						
DB	768	ACGCGCGGAGGCTTCAAGATCGCGCGCATCAGAGGCGGCTTCTTCAAGCTCAACCCCA	827						
QY	1076	TCAACAACAGAGACCCCGCGCATCCGTACCAAGTACAACTGTGTGCGCCCGAGGCTGGAAGG	1135						
DB	828	AGGAGGCGCTCGCCATCTCAACGCGCACGTCGCTGGGCTCCGCGCTCGCGGCCACCGTGA	887						
QY	1136	GCAGCGCGCGCGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCGCTTCCGCGCGCGCA	1195						

DB	888	TGTACGAGCGCAACGTCCTTGGCGGCTCTGTGCGAGGTCCT--GTCCGCGCGCTTCTTCGCG	944
QY	1196	ACCCCGAGATCGTATCTACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1255
DB	945	AGTTCATGACGCGCAACCGCGAGTACAGGACCACTGACCCACAAGCTGAAGACCAACC	1004
QY	1256	AGCACCGCGCGCAAGATGAGGAGCTGCGCAAGCAACCTGCTG--CGCTGGGGCTTACCA	1312
DB	1005	CGGGGTCCATCGAGCG	1064
QY	1313	CCCCGACAGAAGACCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1372
DB	1065	AGCAGGCGCAAGAGTGAACGAGCTGGACCGCGCTGCTGAAGCCCAAGCAGGACAGGTACG	1124
QY	1373	AGTGGACCGTGCAGCG	1432
DB	1125	CGTCCGACGTCGCCCGAGTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1184
QY	1433	AGAAGCTGGTGGCAAGCTGAACCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1492
DB	1185	AGTCCATCGAGCGCGAGGTCAACTCCGTGAACGA---CAACCGGTCTACGACGTCCACC	1241
QY	1493	AGCTGTCAAGCTGCTCG	1552
DB	1242	CGGCAAGCGCGCTGACCG	1301
QY	1553	AGCGCGAGCTGGAGCTGGCGCGAGAACCGGAGATCTTCCGCGAGCGCGCGCGCGCGCG	1612
DB	1302	ACGCGCGCTTCGCCATCGCGCAACATCGGCAAGCTCATGTTGCGCGCAGTTCTCCGAGTCG	1361
QY	1613	ACTACGACCGCGCAAGGACCTGGTGGCGCGAGATCCAGAGCAGGCGCGCGCGCGCGCG	1672
DB	1362	TCAACGAGTTCTACAACACCGCGCTCACTTCAACCTTGGCGCGCGCGCGCGCGCGCGCG	1421
QY	1673	CCTACAGATCTACCAGGACCGCTTCAAGAACCTGAGACCGCGCGCGCGCGCGCGCGCG	1732
DB	1422	TGGACTACGCGCTTCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1481
QY	1733	GCACCG	1792
DB	1482	ACCTGGCAACCG	1541
QY	1793	AGAGCAT---CGTGAATCGGGCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1849
DB	1542	ACTCCCTGGCGCGCTCTCGCGCGAGGAAGACCGCGCGCGCGCGCGCGCGCGCGCGCG	1601
QY	1850	CCTGGGAGACCTGGTGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1909
DB	1602	TGTCGTCACCTACATCGTGGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1661
QY	1910	TGAACACCG	1969
DB	1662	ACATCAAGGCGTGGTGAACACACCGTGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1721
QY	1970	CCGAGACCTTCTACGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	2029
DB	1722	ACCGCTCGGCGAGCTCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1781
QY	2030	ACGTGACCG	2089
DB	1782	ACGCGGAGCGCGCTTACGTACGCGGAGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG	1841
QY	2090	CCGAGTGCAGGCGCATTCAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	2149
DB	1842	AGAAGCTGCG	1901
QY	2150	CCGACAGCG	2209
DB	1902	CCTCCGCTTCTCCAAGATCACCAGGTTTCGAGGAGGAGCTCCGCGCGCGTGTCTGCCCG	1961
QY	2210	TGCTGAA 2216	
DB	1962	AGGTGA 1968	


```
RESULT 2
BM321451
LOCUS BM321451 951 bp mRNA linear EST 03-JAN-2002
DEFINITION rockefeller.0.1211 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC
            3.3.1.1), mRNA sequence.
ACCESSION BM321451
VERSION BM321451.1 GI:18055857
KEYWORDS EST
SOURCE Mastigamoeba balamuthi.
ORGANISM Mastigamoeba balamuthi.
REFERENCE 1 (bases 1 to 951)
AUTHORS Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
        Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
        Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
        Philippe,H.
TITLE The analysis of 100 genes supports the grouping of three highly
JOURNAL divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
COMMENT 21819461
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 951 Std Error: 0.00
POLYA-No.

FEATURES             Location/Qualifiers
     source           1..951
                    /organism="Mastigamoeba balamuthi"
                    /strain="ATCC 30984"
                    /db_xref="taxon:108607"
                    /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                    /notes="syn: Phreatamoeba balamuthi"
BASE COUNT 186 a 321 g 303 t 139 t 2 others
ORIGIN
Query Match          3.5%; Score 85.6; DB 13; Length 951;
Best Local Similarity 46.0%; Pred. No. 0.00011;
Matches 323; Conservative 0; Mismatches 376; Indels 3; Gaps 1;

QY 1722 CGCCAGATGGCGACCGCCACACACAGAGCTGAAGCAGCTGACCGAGCGCGTCAGAA 1781
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 CGCCAGATGGCGTGTGCTGCTGCAACATCTTCTGACGCGAGCACCGCGCGCGC 278

QY 1782 GATCGCATGGAGCATCGTGATCTGGGCAAGACACCCCAAGTTCGCTGCCATCCA 1841
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 CATCGCGAGCGCGCGTCTCGGCTTTCGGCTTGGAGGGCGAGAACCTCCAGGAGTACTG 338

QY 1842 GAAGGAGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTG 1901
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 339 GGAGTGACCTGGAGCGCCTGTCTTCGGCCCTACACGGCCCTCAGATCATCTGTCGA 398

QY 1902 GGAGTTCGTGAACACCCCGCCCTGGTGAAGCTGTGTACCTGAGGAGGAGGCCAT 1961
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 399 CGAGCGCGGTGACGCGACTCTGATGATCCACAGGGGTTTCGGCGCGAGGACACCCCAA 458

QY 1962 CATCGGCGCGAGACCTTCTACGTGGAGCGCGCGCCCAACCGCGAGACCAAGATCGGCA 2021
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 459 GCTGCTGGAGGAGACGAGGCGCTTCGAGGAGGTGCGCTTCAACACGTCTCAAGCA 518

QY 2022 GGCGCGGTAGGTGACCGACCGGGCGCGCAGAGATCGTGAGCTGACCGAGACCCCAA 2081
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 519 GTTCCAGAGGAGACCGCGCTTCTGGACACAGATCTCTCCCGAGATCGCGGTGTAG 578

QY 2082 CCAGAGACCGAGGTGAGCGCATCCAGCTGGCCCTGCAGGACGAGCGGACGAGGTGAA 2141
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 CGAGGAGACGACGACTGGCCGTNATGAGCTGTACCAAGCTGCACGCGACGCGAAGTGT 638

QY 2142 CATCGTGACCGACGACGAGTACGCCCTTGGGCATCATCCAGCGCGACCGCCGACAGACGGA 2201
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 639 GTTCCGCGCGTCAACGTCACGAC---TCTNTACCAAGACAAGTTCGACAACATCTA 695

QY 2202 GAGCGAGCTGGTAACCAAGATCATCGAGCAGCTGATCAAGAAGAGAGAGTGTACTCTGAG 2261
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 696 CGGCTGCGCGCCACTCGCTCATCGAGCGGATCAAGCGCGCACCGACGCTGATGCTCGCGG 755

QY 2262 CTGGTGCGCGCCCAAGAGGCGATCGCGGCAACGAGCAGATCGACAGCTGGTAGCA 2321
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 756 CAAGTTCGCGTCTGTCGCGGCTACGCGACGCTGGGCAAGGGCTGCGCGAGTCTCGG 815

QY 2322 GGGCATCCCAAGGTGCTTCTCTGGAGGCGATCGATGGCGGATCGTGTATACAGTA 2381
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 816 CGGCAGGGCTGCGCGCTCATCGTGACGAGATCGACCCCTCTGCGGCTCGAGCGCTC 875

QY 2382 CATGCAACCACTGTAGTGGGCGAGCGCGCGCTAGGATCGA 2423
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 876 GATGCGCGGCTTCGAGGTTCAACACGCTCGAGGCGGGGCTCGA 917

RESULT 3
BM320864
LOCUS BM320864 1132 bp mRNA linear EST 03-JAN-2002
DEFINITION rockefeller.0.46 Mastigamoeba balamuthi lambda ZAP II Library
            Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
            sequence.
ACCESSION BM320864
VERSION BM320864.1 GI:18055270
KEYWORDS EST
SOURCE Mastigamoeba balamuthi.
ORGANISM Mastigamoeba balamuthi.
REFERENCE 1 (bases 1 to 1132)
AUTHORS Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
        Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
        Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
        Philippe,H.
TITLE The analysis of 100 genes supports the grouping of three highly
JOURNAL divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
COMMENT 21819461
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 1132 Std Error: 0.00
POLYA-No.

FEATURES             Location/Qualifiers
     source           1..1132
                    /organism="Mastigamoeba balamuthi"
                    /strain="ATCC 30984"
                    /db_xref="taxon:108607"
                    /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
                    /notes="syn: Phreatamoeba balamuthi"
BASE COUNT 220 a 413 c 324 g 171 t 4 others
ORIGIN
Query Match          3.3%; Score 80.2; DB 13; Length 1132;
Best Local Similarity 45.3%; Pred. No. 0.00085;
Matches 375; Conservative 0; Mismatches 444; Indels 9; Gaps 2;

QY 139 CGCGCGCGCCCAAGAGGCTGCTGGAAGTGGCGAGGAGGGCCACAGATGAAGAC 198
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 79 CGCGCGGAGGCAAGACGGACTACCGCGCGGCCANCTGGTGATCCAGGACAAGAAC 138

QY 199 TGCACCGAGCGCCAGGCCAACTTCTTCGCGAGGACCTTGGCTTTCGCCAGGGAAGGCC 258
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 139 TACAACAGCCCAAGTACCGCTTCGTC-----GTCCGCTTCCACACAGGACATCGTC 192

QY 259 CGCGAGTTCGCCCGAGCAGACAACCGGCCCAACAGCCCCCAGCGCGGAGCTGCAGGTG 318
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 193 TGCAGATCGCCTACGCCAAGATCGACGGGACCACTCTTCGCGCGCGCTACTCGCAC 252

QY 319 CGCGCGCAACACCCCGCAGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 378
```

```
Db 253 GAGCTCACCCGCTGCGCTGAGCTCGGCTGACCAACTACGCGCGCGCTACGCGACT 312
Qy 379 CAGATACACCTGTGGAGCGCCCTGCTGAGCATCAAGTGGCGGCGGAGCAAGGAG 438
Db 313 GGCTGTGCTGCGCGCCCGCTGCTGAAGAAGCTCAACCTCGACTCCAAGTACGAGGGT 372
Qy 439 GCCTGTGAGACCGCGCGGAGACACGCTGCTGGAGGAGATGAGCCTGCCCGGCAAG 498
Db 373 GTCAAGAAGTCAACGCGGAGGACTACAGTTCGAGAGCTCGACGAGCGGCGCCGCT 432
Qy 499 TGAAGACCCCAAGATGATGGCGGCGATCGGCGGCTTTCATCAAGGTGCGCCAGTAGACAC 558
Db 433 TTCAAGCCCTGCTCGACGCTCGGCGCTGCTCGGCAACCTCGACTGGCGCGCGCTTCGCC 492
Qy 559 ATCCTGATGAGATCTCGGCGAAGAAGCCATCGGACCGCTGCTGATGCGGCCACCC 618
Db 493 GCCCTCAAGGGCATGTGCGACGGGGGCTCAACGTCGCC---CACAGCGAGACCGGCTTC 549
Qy 619 GTGAACATCATCGGCGCAACATGCTGACCCAGCTGGGCTGCACCTGAACTTCCCATC 678
Db 550 GTGCGCTTCAACGCGACAGAAGAGCTCAACGCCGGGTCTCTCCGCAAGTACATCTTC 609
Qy 679 AGCCCCATCGAGACCGTGGCGGTGAAGTGAAGCCCGGCGATGAGCGGCCCAAGGTGAAG 738
Db 610 GCGGCGCACGTCGCGCGGTACATGAAGCTCTCAAGGAGCAGGACGCGCGCTTCGAC 669
Qy 739 CAGTGGCCCTGACCGGAGGAGATCAAGGCCCTGACCGCATCTCGGAGGATGGAG 798
Db 670 CGCCAGTTCGCGCTACGCCAGAGGGTGTACCGCGGACATGCTCGAAGATCTAC 729
Qy 799 AAGGAGGCAAGATCAACAGATCGGCGCGGAGAACCCCTACAAACCCCGCTTTCGCC 858
Db 730 ACCGAGCCACAGAGATCCGCGCGGACCGACCTTCGTCGCCCAAGCGCGCTCGAAG 789
Qy 859 ATCAAGAAGAGGACAGACCAAGTGGCGGAAGCTGGTGGACTTCGCGGAGCTGAACAAG 918
Db 790 CCCGAGGCGGCAAGCCCAAGCACTGGGCGAAGCGGAGGTGAGCTACCAAGGCGCAAG 849
Qy 919 CGCACCCAGGACTTCTGGAGGTGACGTGGGCATCCCAACCCCGC 966
Db 850 AACCGCTGCGCCAGAGAGGTCCGCTGGGCTACCCCGACGCCCC 897
```

RESULT 4
BM320900 1165 bp mRNA linear EST 03-JAN-2002
LOCUS
DEFINITION rockefeller.0.353 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
sequence.
BM320900
BM320900.1 GI:18055306
EST
SOURCE Mastigamoeba balamuthi.
ORGANISM Mastigamoeba balamuthi
Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.

REFERENCE 1 (bases 1 to 1165)
Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.
The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
21819461
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockefeller.edu
Insert Length: 1165 Std Error: 0.00
POLYA-No.

TITLE The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE 21819461
COMMENT Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockefeller.edu
Insert Length: 1165 Std Error: 0.00
POLYA-No.

FEATURES
source
1. .1165

```
/organism="Mastigamoeba balamuthi"  
/strain="ATCC 30984"  
/db_xref="taxon:108607"  
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
/note="syn: Phreatamoeba balamuthi"  
BASE COUNT 221 a 426 c 336 g 178 t 4 others  
ORIGIN  
Query Match 3.2%; Score 79.6; DB 13; Length 1165;  
Best Local Similarity 45.1%; Pred. No. 0.0011;  
Matches 428; Conservative 0; Mismatches 506; Indels 16; Gaps 3;  
Qy 90 CATCAAGTGTTCACCTGCGCAAGAGGGCCACATCGCCGCAACTGCGCGCCCGCCCG 149  
Db 46 CGTCAAGAACAAAGCGTACTTCAAGCGCTTCCAGACCAAGTTCGCTGCGCGCGAGG 105  
Qy 150 CAAGAAGGGTGTGGAAGTCCGCAAGAGGGCCACCAGATGAAGGACTGCACCGAGCG 209  
Db 106 CAAGACGGACTTACCGCGCGGCCANCTGGTGATCCAGACAAAGAACAGTACACAGCCC 165  
Qy 210 CAGGCGCAACTTCTTCCGCGAGGACCTGGCCTTCCCGCAGGCAAGGCCCGCGAGTTCCC 269  
Db 166 CAAGTACCGCTTCGTC-----GTCCGCTTCAACAAGGACATCGTCTCCAGATCGC 219  
Qy 270 CAGCGAGAGAACCGCGCAACAGCCCGCACAGCGCGGAGCTGCAAGTGGCGGCGACAA 329  
Db 220 CTACGCGAAGATCAGCGCGACACATCTCGCGCGCGCTTACTCGCACGAGCTCACCCG 279  
Qy 330 CCCCCGAGGAGCGCGCGCGCGAGCGCGGCGACCTGAACCTTCCCCAGATCACCCCT 389  
Db 280 CTTGGCGGTCAAGTTCGCGCTGACCAACTACGCGCGCGCTTACCGGACTGGCCTGCTGCT 339  
Qy 390 GTGCAGCGCCCTGTGTGAGCATCAAGGTGGCGGCGCAGATCAAGGAGGCCCTGCTGGA 449  
Db 340 GCGCGCGCTGTGTGAAGACTCAACCTCTACTCAAGTACGAGGGTGTCAAGAAGGT 399  
Qy 450 CACCGCGCGCGACGACACCGTGTGTGAGGAGATGAGCCTGCCCGGCAAGTGGAGGCCAA 509  
Db 400 CAAGCGGAGGACTACAACTCGAGAGCTCGAGGAGCTCGAGGAGCGCGCGCTTCAAGGCC-- 457  
Qy 510 GATGATCGCGGCATCGCGGCTTCTATCAAGGTGCGCCAGTACGACACGATCTCTGATCGA 569  
Db 458 --TGCTGACGTGCGGCTGTGCGGACCTGACTGGCGCGCGCTGTTCGCGCGCCCTCAA 515  
Qy 570 GATCTGCGCAAGAGGCGCATCGGCACCGTGTGATCGCGCGCCCGCCCGCTGAACATCAT 629  
Db 516 GGGCATGTGCGACGCGCGCGTCAAGCTCCCGCACAGCAGAGACCCGCTTCGTGGCTTCAA 575  
Qy 630 CGCGCGCAACATGCTGACCGAGCTGGGCTGCACCTGAACTTCCCATCAGCGCCCATCGA 689  
Db 576 CGGCGCAAGAGGAGCTCAACGCGCGGTCTCCGCAAGTACATCTTTCGCGCGCACGT 635  
Qy 690 GACGTGCGCGTGAAGCTGAAGCGCGCATGGAGGCGCGCGCGCGCGCGCGCGCT 749  
Db 636 CGCGCGTACATGAAGTCTCTCAAGGACGAGACGCGCGCGCTTCGACCGCGCTTCTC 695  
Qy 750 GACCGAGGAGAGATCAAGGCGCTGACCGCCCTCTGCGGAGGATGGAGAAGGGGCAA 809  
Db 696 CGCTACCCCAAGGAGGGTGTCAACGCGGACATGCTCGAGAAGATCTACACGAGGGCCCA 755  
Qy 810 GATCACCAGATGCGCGCGCGAGACCCCTACACACCCCGCTGTTCGCCATCAAGAAGAA 869  
Db 756 -----CAAGCAGATTCGCGCGCGCGCGCGCGCGCGCGCGCGCTCGAAGCCCGCA 809  
Qy 870 GGACGACCAAGTGGCGCAAGCTGGTGAGCTTCCGCGAGCTGAACAAAGCGCACCCAGGA 929  
Db 810 GGGCGCCCAAGCCCAAGCACTGGGCGAAGCGGAGGTGACGTACCAAGAGCGCAAGACCG 869  
Qy 930 CTTCTGGGAGGTGAGCTGGGCATCCCCACCCCGCGCGCTTGAAGAAGAAGAGAGCGT 989  
Db 870 CGTCCGCCAAGAAGAGGTCCCGCTGGGCTACCCCGACGCGCGCGCGCGCGCTTAAATCC 929  
Qy 990 GACCGTGTGGAGTGGCGGAGCGCTACTTTCAGCGTGGCGCGCGCGCGCGCGG 1039
```

Db 930 CGGTGCCCTGTACACCGCCATCTCCCGTGGCGGTTCGGTGTGNGCG 979

RESULT 5

BM321430 867 bp mRNA linear EST 03-JAN-2002
LOCUS rockefeller.0.1153 Mastigamoeba balamuthi lambda ZAP II Library
DEFINITION Mastigamoeba balamuthi cDNA similar to ribosomal protein S4, mRNA

sequence.
BM321430

ACCESSION BM321430.1 GI:18055836

VERSION BM321430.1

KEYWORDS EST.

SOURCE Mastigamoeba balamuthi.

ORGANISM Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.

REFERENCE 1 (bases 1 to 867)

AUTHORS Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,

Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and

Philippe,H.

TITLE The analysis of 100 genes supports the grouping of three highly

divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

21819461

CONTACT: Muller Miklos

Laboratory of Biochemical Parasitology

The Rockefeller University

1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu

Insert Length: 867 Std Error: 0.00

POLYA=No.

FEATURES

source Location/Qualifiers

1..867

/organism="Mastigamoeba balamuthi"

/strain="ATCC 30984"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/notes="syn: Phreatamoeba balamuthi"

BASE COUNT 185 a 298 c 254 g 124 t 6 others

ORIGIN

Query Match 3.2%; Score 79; DB 13; Length 867;

Best Local Similarity 45.7%; Pred. No. 0.0013;

Matches 314; Conservative 0; Mismatches 370; Indels 3; Gaps 1;

QY 58 CAGCGCAGCACTTCAAGGCCCCAAGCGCATCATCAAGTCTTCAACTCGGCAAGGAG 117

Db 131 CCGCACAGATGCGGAGTGCTTCCCGGTATCATCTCTGTGCGCAACAGGTTGAAGTAC 190

QY 118 GGCACATCGCCCGCAACTGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGGAAG 177

Db 191 GCGCTGACCCCGCTGAGGTCACTTCATCGTGATGAGCGCTGATCAAGATCGACGGC 250

QY 178 GAGGGCCACCATGAGAGACTGACCGAGCGCCAGCGCAACTTCTTCGCGGAGGACTG 237

Db 251 AAGTCCGCGCACCGACACGCGACTTCCCGCGGGCTTTCATGAGACGCTCTCGATCGACAAG 310

QY 238 GCCTTCCCGCAGGCGAGCGCGAGTTCGCCAGCAGGACGAGCGCGCCCAAGCCCC 297

Db 311 ACCGACGAGCACTTCCGCTGCTTACGACACCAAGGCGCTTCCAGGCGCACCGCATC 370

QY 298 ACCAGCGCGAGCTGCAAGT---GCGCGGGGACAAACCCCGCAGCGAGCGCGCGCGGAG 354

Db 371 AACTCGACCGAGGCGCAAGTCAAGCTCGGCAAGTCCGCGCGGTGCGAGTCTGGCAACAG 430

QY 355 GCGCAGGCGACCTTGAACTTCCCGCCAGATCACCTGTGGCAGCGCCCGCTTGGTGAGATC 414

Db 431 GGCATCCCGTACCTGTGTGAGCGCGACGACGCGCGACGATCCGCTACCCCAACCCCGACATC 490

QY 415 AAGTGTGGCGCCGATCAAGGAGGCGCTGCTGACACCGCGCGCGACACCGTCTGCT 474

Db 491 AAGTCAACGACAGGTCGAAGATCGACCTGGCTCGGGCAAGATCATCGACTTCGTGAGG 550

QY 475 GAGGAGATGAGCTGCGCGCAAGTGGAGCCCAAGATGATCGGCGCATCGCGGCTTC 534

Db 551 TTCGAGATCGGCAACCTCGTATGATCACTGGCGGACCACTTGGCGGCTCGGCGTC 610

QY 535 ATCAAGTGGCCAGTAGCAGCAGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGC 594

Db 611 ATTGTGCGCGCAGAGAACGAGGCGTCTCGAGATCATCCAGTCAAGGACCGCGTC 670

QY 595 ACCGTGCTGATGCGCGCCACCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTG 654

Db 671 GGCCACCAAGTTCGCGACGCGCTGACCAACGCTTCGTGATCGGCAAGGCGCAAGTCC 730

QY 655 GGTGTCACCTGAACTTCCCATCATGCGCCATCGAGACCGTGCCTGAACTGAAGTCC 714

Db 731 CTCGTACGCTGCGCGCGGCAAGGCAATCAAGAGTGCATCATCGAGGATTCAGGCG 790

QY 715 GGCATGAGCGCCCAAGGTGAAGCAG 741

Db 791 CGCCACGCGCACAGGACCGCAGCAGGAG 817

RESULT 6

BM321022 1550 bp mRNA linear EST 03-JAN-2002

LOCUS rockefeller.0.1192 Mastigamoeba balamuthi lambda ZAP II Library

DEFINITION Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC

3.3.1), mRNA sequence.

ACCESSION BM321022

VERSION BM321022.1 GI:18055428

KEYWORDS EST.

SOURCE Mastigamoeba balamuthi.

ORGANISM Mastigamoeba balamuthi.

REFERENCE 1 (bases 1 to 1550)

AUTHORS Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.

Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,

Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and

Philippe,H.

TITLE The analysis of 100 genes supports the grouping of three highly

divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

21819461

CONTACT: Muller Miklos

Laboratory of Biochemical Parasitology

The Rockefeller University

1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu

Insert Length: 1550 Std Error: 0.00

POLYA=No.

FEATURES

source Location/Qualifiers

1..1550

/organism="Mastigamoeba balamuthi"

/strain="ATCC 30984"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/notes="syn: Phreatamoeba balamuthi"

BASE COUNT 282 a 528 c 495 g 226 t 19 others

ORIGIN

Query Match 3.1%; Score 76.6; DB 13; Length 1550;

Best Local Similarity 45.6%; Pred. No. 0.0034;

Matches 308; Conservative 0; Mismatches 364; Indels 3; Gaps 1;

QY 1749 CGAGCTGAACAGCTGACCGAGCGGTGCAGAAAGATCCCATGCGAGACATCGTGATCTG 1808

Db 5 CGGCACGAGGACGAGGACCAACGCGCGCGCGCATCGCGAGCGCGCTCTCGGTCTT 64

QY 1809 GGGCAAGACCCCAAGTTCGCTTCCCATCCAGAAAGAGACCTGGGAGACCTGGTGGAC 1868

Db 65 CGCTTGAAGGGCGAGAACCTCCAGGAGTACTGGAGTGCACCTGGGAAGGCCCTGTGCTT 124

QY 1869 CGATCTAGGACGCCACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCGCTGCT 1928

Db 125 CGGCCCCCTACCAGGGCCCTCAGATCATCGTCGACGACGCGCGTGAAGGCGACTCTGATGAT 184


```

Db 596 GAGGGCCCAAGCCCAAGCACTGGGGCAAGCGCAGGCTGACGTACCAAGCCGCAAGAAC 655
QY 928 GACTTCTGGGAGGTGCGAGCTGGGCGATCCCGACCCCGCGCTGAAGAAGAAAGAGC 987
Db 656 CGCGTCCCGCAGAGAGGTCCGCTGGGTACCCGACGCCCCCAAGACCGATTAATT 715
QY 988 GTGACCGTGTGACGCTGGGCGAGCGCTACTTCAGCG 1024
Db 716 CCGGTGCGCTGTACACGCGCATCTCCCGCTGGCGG 752

RESULT 10
BM587428/c
LOCUS
DEFINITION
  BM587428 566 bp mRNA linear EST 25-FEB-2002
  19600449696110 5', mRNA sequence.
ACCESSION
  BM587428
VERSION
  BM587428.1 GI:18883289
KEYWORDS
  EST.
SOURCE
  African malaria mosquito.
ORGANISM
  Anopheles gambiae
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
  Anopheles.
REFERENCE
  1 (bases 1 to 566)
  Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
  ,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
  Celera Anopheles gambiae EST project
  Unpublished (2002)
  Contact: Holt R.A.
  Celera Genomics
  45 W. Gude Dr.,
  Rockville, MD 20850, USA
  Tel: 2404533151
  Fax: 2404534580
  Email: Holt@RAcelera.com
  Plate: NU01004AAX row: G column: 08
  Seq primer: M13 Reverse.
FEATURES
  source
  1..566
    /organism="Anopheles gambiae"
    /strain="RSP-ST (Reduced susc. to Permethrin - std.
    chromosome)"
    /db_xref="taxon:7165"
    /clone="19600449696110"
    /clone_lib="A.Gam.ad.cDNA.bloodi"
    /dev_stage="Adult"
    /lab_host="DH10b"
    /note="Vector: pSport1; Site_1: SalI; Site_2: NotI; Whole
    adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
    hours after human blood feeding. cDNA inserts >500 bp
    cloned directionally into pSport 1. Not 1 site is 3'.
    Clones available through the Malaria Research and
    Reference Reagent Resource Center (www.malaria.mr4.org)"
  BASE COUNT 60 a 202 c 165 g 139 t
  ORIGIN
    Query Match 3.0%; Score 72.8; DB 13; Length 566;
    Best Local Similarity 48.3%; Pred. No. 0.014;
    Matches 203; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

QY 1944 GCTGGAGAGGAGCCCATCATCGCGCGCAGAGACTTCTAGCTGGACGGCGCCCAACCG 2003
Db 482 GTTCGAGATCGAGCCCAAGGCGATCTCGAGGTGTGCGCGAGGCAAGGCGCGGCAA 423
QY 2004 CGAGACCAAGATCGGAGCGCGCTACGTACCGACCGCGGGCGCGAGAGATCGTGAG 2063
Db 422 CCGGAAAAGATCGTCATCACCACAGCAGCAACCGCCCTGACCCCGGAGACATCGAGCG 363
QY 2064 CTTGACCGGAGACCAACCAAGAGACCGAGCTGCGAGGCATCCAGCTGGCCCTGCAGGA 2123
Db 362 CATGATCAAGGATCGGAGCGGTTCGCCCGACGACGACAGAGCTGAAGGAGCGCGTGA 303
QY 2124 CAGCGGAGGAGGTGAATCTGTGACCGAGACCCAGTACGCGCTGGGCATCATCCAGGC 2183

```

```

Db 302 GGCCCGCAACGAGCTCGAGAGCTACGCGTGAAGAACCCAGCTCAGCTCGAAGGA 243
QY 2184 CCAGCCCGCAAGAGCGCAGAGCGAGCTGGTGAACCAACATATCGAGCAGCTGATCAAGAA 2243
Db 242 CAAGCTGGGCGCGAGCGGTGTCGCGACGACGACAAAGCCAAAGATGGAGGCGGATCGACGA 183
QY 2244 GGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCAAGGCGCATCGCGGCAACGAGCAGAT 2303
Db 182 GAAGATCAAGTGGCTGCGAGAGAACGAGGACCGAGCGGAGAGATCAAGAAGCAGAA 123
QY 2304 CGACAAGCTGGTGAGCAAGGCGCATCCGCAAGGTGCTTCTCGTCGACGGCGATCGATGCGG 2363
Db 122 GAAGGAGCTGGAGGACATCGTGCGAGCCCATTCATTGCCAAGCTGTACGCGAGCAGTGGCGG 63

RESULT 11
BE601575
LOCUS
DEFINITION
  BE601575 640 bp mRNA linear EST 22-OCT-2001
  HVSMEH0098K08f Hordeum vulgare 5-45 DAP spike EST library
  HVCDNA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEH0098K08f,
  mRNA sequence.
ACCESSION
  BE601575
VERSION
  BE601575.3 GI:16322423
KEYWORDS
  EST.
SOURCE
  Hordeum vulgare.
ORGANISM
  Hordeum vulgare
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
  ; Triticeae; Hordeum.
  1 (bases 1 to 640)
REFERENCE
  Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
  ,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton
  ,R.D., Close,S.J., Oates,R. and Main,D.
  Development of a genetically and physically anchored EST resource
  for barley genomics: Morex 5-45 DAP spike cDNA library
  Unpublished (2001)
  On Aug 21, 2000 this sequence version replaced gi:13190104.
  Contact: Wing RA
  Clemson University Genomics Institute
  Clemson University
  100 Jordan Hall, Clemson, SC 29634, USA
  Tel: 864 656 7288
  Fax: 864 656 4293
  Email: rwing@clemson.edu
  Total hg bases = 461
  Seq primer: AATTACCTCTACTAAAGG
  High quality sequence stop: 534.
  Location/Qualifiers
    1..640
      /organism="Hordeum vulgare"
      /cultivar="Morex"
      /db_xref="taxon:4513"
      /clone="HVSMEH0098K08f"
      /clone_lib="Hordeum vulgare 5-45 DAP"
      /tissue_type="5-45 DAP Spike"
      /lab_host="SOLR"
      /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
      Plants were grown in the greenhouse at the University of
      California, Riverside (Fenton, SJ Close, TJ Close). Whole
      spikes with awns trimmed were collected at 5, 10, 15, 20,
      30 and 45 DAP (Fenton). Total RNA was prepared from each
      pool, equal quantities of all six RNA pools were combined,
      poly(A) RNA was purified from the mixture, one primary
      unamplified cDNA library was made, and 1 million pfu were
      in vivo excised to give pluescript SK(-) cDNA phagemids
      (Choi) in the TJ Close lab at the University of California,
      Riverside. Phagemids were plated and picked at the Clemson
      University Genomics Institute (CUGI) (Begum, Palmer,
      Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
      sequencing and sequence analysis were performed at CUGI
      (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).

```

The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see

<http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinjohs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)

```

BASE COUNT      135 a      230 c      177 g      97 t
ORIGIN

Query Match      2.9%; Score 72.2; DB 10; Length 640;
Best Local Similarity 47.7%; Pred. No. 0.017;
Matches 209; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 789 GGAGATGGAGAGAGGGCAAGATCACAGATCGGCGCCGAGAACCCCTACAAACACCCC 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110 GGAGAGGTGACGGTGAAGGTGTGCGGCAAGATGATCTCGGTGACGGGCGCGCGCAC 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 849 CGTGTTCGCCATCAAGAGAGGACACCAAGTGGCGGCAAGCTGTGGACTTCCCGGA 908
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170 CTTGACCGCAACTTCAAGCACCTCAACCTCGACTTCCAGCTGACGAGCGGCGGCGAA 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 909 GCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCATCCGCCACCCCGCGG 968
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 230 GCTCAAGGTGACCGCTGTGGCACCAGCGCCGACCATGCGCCCATCCGCCCGCAT 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 969 CTTGAAGAAGAAGAGCGTGACCGGTGCTGGAGCTGGGCGACCGCTTCTACAGCTGCC 1028
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 CTCCACGTCAGAACCTCATACCGGCGTCAACCAAGGGCTTCGGCTTACAAGATGCGCTT 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1029 CTTGAGAGGACTTCCGCAAGTACCGGCTTCAACATCCCGAGATCAACACAGAGAC 1088
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 CGTCTAGCTCATTCCCATCAACGCGCTCATACCGCGCCGCAACCGGCGCATCGAGAT 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1089 CCGCGCATCGGTACCAAGTACAGCTGCTGCCCGAGGGCTGGAAGGCGCCCGCAGAT 1148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 CCGCAACTTCTTCGGCGAGAGAGGTGAGGAGGTGGACATGCTCAGCGGGTCAACAT 469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1149 CTTCCAGAGCATGATACCAAGATCTTGGAGCCCTTCGCGCGCGCAACCCCGAGATCGT 1208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 CTTGCGGTCCGAGAAGGTCAAGATGATGATCGTCTCTCGAGCGCAACGACATCGAGCTCGT 529
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1209 GATCTACAGGCCCGCCCT 1226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 530 NTCCGCTCCGCGCCCT 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 12
BM372120
LOCUS
DEFINITION
  BM372120 Ebro03_SQ004_B09_R root, 3 week, waterlogged, cv Optic, Ebro03
  Hordeum vulgare cDNA clone Ebro03_SQ004_B09 5', mRNA sequence.
ACCESSION
  BM372120
VERSION
  BM372120.2 GI:21948499
KEYWORDS
  EST.
SOURCE
  Hordeum vulgare.
  Hordeum vulgare
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
  ; Triticeae; Hordeum.

```

```

REFERENCE
  1 (bases 1 to 500)
  Hedley,P., Liu,H., Caldwell,D., McCallum,N., Muddle,S., Cardle,L.,
  Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
  Development of Barley Transcriptome Resources
  Unpublished (2001)
  On Jan 10, 2002 this sequence version replaced gi:18115510.
  Contact: Waugh R, Marshall DF
  Genome Dynamics/Computational Biology
  Scottish Crop Research Institute
  Invergowrie, Dundee, DD2 5DA, Scotland, UK

```

Tel: 00 44 1382 562731
 Fax: 00 44 1382 562426
 Email: est@scri.sari.ac.uk
 All sequence has a Phred quality score of 20 or over
 Seq primer: M13 reverse.

FEATURES

```

source
  1..500
  /organism="Hordeum vulgare"
  /cultivar="Optic"
  /db_xref="taxon:4513"
  /clone="Ebro03_SQ004_B09"
  /clone_lib="root, 3 week, waterlogged, cv Optic, Ebro03"
  /tissue_type="root"
  /dev_stage="3 week"
  /lab_host="DH10B"
  /note="vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
  Non-normalised library, directionally cloned into pSPORT1.
  Derived from roots of 3 week old waterlogged barley
  plants. Developed as part of the barley transcriptome
  resources of BBSRC/SEERAD funded cereal IGF (Investigating
  Gene Function) project."
  Gene Function
BASE COUNT      108 a      177 c      139 g      76 t
ORIGIN

```

```

Query Match      2.9%; Score 71.6; DB 13; Length 500;
Best Local Similarity 47.7%; Pred. No. 0.021;
Matches 209; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 789 GGAGATGGAGAGAGGGCAAGATCACCAAGATCGGCGCCGAGAACCCCTACAAACACCCC 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 GGAGAGGTGACGGTGAAGGTGTGCGGCAAGATGATCTCGGTGACGGGCGCGCGCAC 91
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 849 CGTGTTCGCCATCAAGAGAGGACACCAAGTGGCGCAAGCTGTGGACTTCCCGGA 908
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92 CTTGACCGCAACTTCAAGCACCTCAACCTCGACTTCCAGCTGACGAGCGGCGCGCAA 151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 909 GCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCATCCGCCACCCCGCGG 968
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 GCTCAAGGTGACCGCTGTGGCACCAGCGCGCACCATGCGCGCATCCGCCCGCAT 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 969 CTTGAAGAAGAAGAGCGGTGACCGTCTGGAGCTGGGCGACCGCTTCTTACAGCTGCC 1028
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 CTCCACGTCAGAACCTCATACCGGCGTCAACCAAGGGCTTCCGCTACAAGATGCGCTT 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1029 CTTGAGAGGACTTCCGCAAGTACACCGCTTCAACATCCCGAGATCAACACAGAGAC 1088
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 CGTCTAGCTCATTCCCATCAACGCGCTCATACCGCGCCGCAACCGGCGCATCGAGAT 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1089 CCGCGCATCGGTACCAAGTACAACTGCTGCCCGAGGGCTGGAAGGCGACCCCGCAGAT 1148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 332 CCGCAACTTCTTCGGCGAGAGAGGTGAGGAGGTGGACATGCTCAGCGGGTCAACAT 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1149 CTTCCAGAGCATGATACCAAGATCTTGGAGCCCTTCGCGCGCGCAACCCCGAGATCGT 1208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 392 CTTGCGGTCCGAGAAGGTCAAGATGATGATGATCGTCTCTCGAGCGCAACGACATCGAGCTCGT 451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1209 GATCTACAGGCCCGCCCT 1226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 CTCCGCTCCGCGCCCT 469
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 13
BM368580
LOCUS
DEFINITION
  BM368580 Ebro03_SQ004_A06_R embryo, 40 DPA, no treatment, cv Optic, Ebro08
  Hordeum vulgare cDNA clone Ebro08_SQ004_A06 5', mRNA sequence.
ACCESSION
  BM368580
VERSION
  BM368580.2 GI:21937722
KEYWORDS
  EST.
SOURCE
  Hordeum vulgare.
  Hordeum vulgare
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:14:58 ; Search time 79.8051 Seconds
(without alignments)
9441.811 Million cell updates/sec

Title: US-09-610-313-32
Perfect score: 2457
Sequence: 1 gtcagccaccatggccga.....gggctagcaccgtgaattc 2457

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15333831 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCRUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1081.6	44.0	2601	4	US-09-117-217-7
2	1081.6	44.0	2601	4	US-09-117-217-9
3	1081.6	44.0	2601	4	US-09-117-217-11
4	1081.6	44.0	2601	4	US-09-117-217-13
5	1075.2	43.8	7399	2	US-08-418-848A-9
6	1075.2	43.8	9709	2	US-08-188-583-5
7	1075.2	43.8	9709	3	US-08-388-353-1
8	1075.2	43.8	9709	3	US-08-488-551B-1
9	1075.2	43.8	9709	4	US-09-309-572-15
10	1075.2	43.8	12494	4	US-08-935-312-13
11	1075.2	43.8	12494	4	US-08-848-760B-33
12	1075.2	43.8	15581	3	US-08-646-538-35
13	1075.2	43.8	15581	4	US-09-503-222-35
14	1065.6	43.4	9737	2	US-08-944-449-7
15	1065.6	43.4	9737	4	US-09-353-362-7
16	1030.4	41.9	9746	1	US-08-022-835-3
17	1030.4	41.9	9746	2	US-08-388-809-3
18	1030.4	41.9	9746	2	US-08-647-714-3
19	1029.2	41.9	8933	4	US-09-124-900-1
20	1029.2	41.9	8933	3	US-08-463-210-4
21	1029.2	41.9	8933	4	US-09-620-958A-3
22	1029.2	41.9	8933	4	US-09-620-958A-9
23	1029.2	41.9	8933	4	US-09-620-958A-9
24	1016.8	41.4	9739	1	US-08-022-835-1
25	1016.8	41.4	9739	1	US-08-388-809-1
26	1016.8	41.4	9739	1	US-08-647-714-1
27	1011.6	41.2	5362	3	US-08-463-210-5

28	999.8	40.7	3033	2	US-07-743-357-19	Sequence 19, Appl
29	999.8	40.7	4113	2	US-07-743-357-21	Sequence 21, Appl
30	994	40.5	9207	3	US-08-388-353-800	Sequence 800, App
31	994	40.5	9207	3	US-08-488-551B-800	Sequence 800, App
32	904.2	36.8	3856	2	US-07-743-357-20	Sequence 20, Appl
33	898.8	36.6	2739	2	US-07-743-357-16	Sequence 16, Appl
34	818.4	33.3	2348	2	US-08-876-546A-1	Sequence 1, Appl
35	818.4	33.3	2348	4	US-09-412-252-1	Sequence 1, Appl
36	818.4	33.3	2348	4	US-09-079-675-1	Sequence 1, Appl
37	805.8	32.8	9793	1	US-08-470-202-56	Sequence 56, Appl
38	805.8	32.8	9793	1	US-08-471-770-56	Sequence 56, Appl
39	805.8	32.8	9793	2	US-08-468-059-56	Sequence 56, Appl
40	805.8	32.8	9793	4	US-09-109-916-56	Sequence 56, Appl
41	798.2	32.5	1680	5	PCT-US93-06748-1	Sequence 1, Appl
42	578.8	23.6	1256	4	US-09-158-695-18	Sequence 18, Appl
43	527.8	21.5	3168	2	US-08-659-251-14	Sequence 14, Appl
44	527.8	21.5	3168	4	US-09-256-490-14	Sequence 14, Appl
45	527.8	21.5	3168	5	PCT-US96-11445-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-117-217-7
; Sequence 7, Application US/09117217
; Patent No. 6221578
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117,217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(492)
; OTHER INFORMATION: gag Polyprotein
US-09-117-217-7

Query Match	44.0%	Score 1081.6;	DB 4;	Length 2601;
Best Local Similarity	67.9%	Pred. No. 4e-169;		
Matches 161;	Conservative	0;	Mismatches 734;	Indels 28; Gaps 6;
QY	14	TGCGCGGGCCATGAGCGAGCCACCA---CGCCCAACATCCTGATGCGCGCAGCAACT	70	
Db	77	TGCGTGAAGCAATGAGCGAGCCACCAATTAACAATTCAGTACCATAATGATGCGAGAGGCATT	136	
QY	71	TCRAGGCGCCCAACCGCATCATCAAGTCTTCAACTGGCGCAAGAGGGCCACATCGGCC	130	
Db	137	TTAGGAACCAAGAAAGATTGTTAAGTGTTCATTTGTCGCAAGAGGGCCACACAGCA	196	
QY	131	GCAACTGCGCGCGCCCGCCGCAAGAGGGCTCTGGAAGTGGCGCAAGAGGGCCACCA	190	
Db	197	GAATTCAGAGGGCCCCCTAGGAAAAGGGCTTTGGAATGTGGAAGAGGAGGACCAAAA	256	
QY	191	TGAAGGACTGCACCGAGCGCCGCAACTTCTTCGCGAGGAGCTGCTTCCCGCCAGG	250	
Db	257	TGAAAGATTGTACTGAGAGACAGCGCTTA-TTTTITGGGAGAGATCTGCGCTTCTTACAG	315	
QY	251	GCAAGGCGCGGAGTTTCCCGAGCGAGAGAACCGCGGCAACAGCCCCACAGCGCGAGC	310	
Db	316	GGAAGCGCGGGAATTTTCTTTCAGAGCAGACAGCCACAGCCACAGGAGAGAGC	375	

1876 CCACAGAAAGCATAGTAATAATGGGAAAGACTCTTAATTTAAACTGCCCCATACAAAGG 1935
1847 AGACCTGGGAGACCTGGTGACCACTACTGCGAGGCCACCTCGATCCCGCAGTGGAGT 1906
1936 AATCATGGGAAACATGGTGACAGATATTGGCAAGCCACCTCGATTCCTTGAGTGGAGT 1995
1907 TCCTGAACACCCCCCTCGTGAAGCTGTGGTACCAGCTGGAGAGAGGCCCATCATCG 1966
1996 TTGTTAATACCCCTCCCTTAGTGAATATTGGTACCAGTTAGAGAAAGACCCATAGTAG 2055
1967 GCGCCGAGACCTTCTACGTGGAGCGGCCGCCCAACCGCGAGACCAAGATCGGCAAGCCG 2026
2056 GAGCAGAAACCTTCTATGTAGATGGGCGACCTAACAGGGAGACTAAATTAGGAAAGCAG 2115
2027 GCTACGTGACCGAGCGGGCGCGCAGAGATCGTGAGCCCTGACGAGACCAACACAGAG 2086
2116 GATATGTTACTAATAGAGGAAGCAAAAGTTGTCAACCTTAAGTACACAAACAAATCAGA 2175
2087 AGACCGAGCTGCAGGCGATCCAGCTGGCCCTGCAGGACAGCGGAGGTGAACATCG 2146
2176 AGACTGAGTTACAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAGAGTAAACATAG 2235
2147 TGACCGACAGCAGTAGTACGCCCTGGGCATCATCCAGGCCCGCAGCCGACAGAGCGAGCG 2206
2236 TAACAGACTCACAAATATGCATTAGGAATCATTAAGCACAACCCAGATCAAAGTGAATCAG 2295
2207 AGCTGTGAACACAGATCATCGACAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 2266
2296 AGTTAGTCAATCAATTAATAGACAGTAAATAAAGGAAAGGTTCTATCTGGCATGGG 2355
2267 TGCCCGCCCAACAGGGCATCGCGCGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGGCA 2326
2356 TACCAGCACAAAGGAATTTGGAGGAATGAACAAGTAGATAATAGTCAGTGTCTGGAA 2415
2327 TCOCGAAGTGCTGCTTCCTGGAGCGGCATCGAT 2358
2416 TCAGAAAGTACTATTTTATAGATGGAATAGAT 2447

RESULT 3

US-09-117-217-11
; Sequence 11, Application US/09117217
; Patent No. 6221578
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kuit
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117,217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (453)...(749)
; OTHER INFORMATION: Protease
US-09-117-217-11

Query Match 44.0%; Score 1081.6; DB 4; Length 2601;
Best Local Similarity 67.9%; Pred. No. 4e-169;
Matches 1610; Conservative 0; Mismatches 734; Indels 28; Gaps 6;
QY 14 TGCCGAGGCCATGAGCCAGGCCACCA---GCGCCAAACATCTGTATGCGAGCGCAGCAACT 70
DB 77 TGGCTGAAGCAATGAGCCCAAGTAACAAATTCAGCTACCAATAATGATGCAGAGAGGCAATT 136

QY 71 TCAGGSCCCCAAGCGCATCATCAAGTGTCTCAACTGCGCAAGGAGGCCACATCGCCC 130
DB 137 TTAGGAACCAAAAGAAAGATTGTTAAGTGTTCATTTGCGAAAGAGGCGACACAGCCA 196
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGCAGGAGAGGGCCACACAGA 190
DB 197 GAAATTGCGAGGGCCCTAGCAAAAAGGGCTGTGGAAATGTGAAAGAGGAAGCACACAAA 256
QY 191 TGAAGGACTCAGCGAGCGCCAGGCCCAACTTCTCCCGGAGGACCTGGCCCTTCCCCCAGG 250
DB 257 TGAAGATTTACTGAGAGACAGGCTAA-TTTTTTAGGGAAGATCTGGCCCTTCTCTACAAG 315
QY 251 GCAAGGCGCGCGAGTTCGCCAGCGAGCAGAAACCGCGCAACAGCCCGCCAGCGCGCAGC 310
DB 316 GGAAGCGCAGGGAATTTTCTCAGAGCAGACAGAGCCACAGCCCGCCAGAGAGAGC 375
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGAGCGCGCAGGGCA 364
DB 376 TTCAGGTTCTGGGTAGAGACAACAACCTCCCTCAGAAAGCAGGAGCGGATAGACAAGAA 435
QY 365 -----CCCTGAACTTCCCGCCAGATCACTCTGTGGCAGCGCCCGCTGTGAGCATCAAG 418
DB 436 CTGTATCTCTTAACTTCCCTCAGGTCACTCTTTGGCAACGACCCCTCGTCAACATAAGA 495
QY 419 TGGCGCGCCAGATCAAGGAGGCGCTGCTGCACACCGCGCGCAGCACCGTGTCTGAGG 478
DB 496 TAGGGGGCAACTAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 555
QY 479 AGATGAGCCTTCCCGCCAGACTGGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
DB 556 AATATGAGTTTGCAGGAAGATGGAACCAAAATGATAGGGGAATTTGAGGTTTTATCA 615
QY 539 AGTGCGCAGTAGGACCAGATCCTGATCGAGATCTCGCGCAAGAGGCCATCGGCACCG 598
DB 616 AAGTAAGACAGATGATGATCAGATCTCATAGAAATCTGTGGACATATAAAGCTATAGGTACAG 675
QY 599 TGCTGATCGGCGCCCGCCCGTGAACATCATCGCGCCCAACATGCTCAGCCAGCTGGCT 658
DB 676 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTGACTCAGATTGGTT 735
QY 659 GCACCTGAACTTCCCGCATCAGCCCCATCGAGACCGGTGCGCGTGAAGCTGAAGCCCGCA 718
DB 736 GCACCTTAAATTTTCCCATTTAGCCCTATTGAGACTGTACCAGTAAATTTAAAGCCAGAA 795
QY 719 TGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGCCCTGACCG 778
DB 796 TGGATGCGCCCAAGGTAAACAATGGCCATTGACAGAAGAAAAATAAAGCATTAGTAG 855
QY 779 CCATCTGCGAGGAGATGGAGAAGGCGCAAGATCACCAAGATCGGCCCGCAGAACCCCT 838
DB 856 AATTTGTACAGAGATGGAAGAGAGAGGAGGAAATTTCAAAATTTGGCCCTGAAATCCAT 915
QY 839 ACAACACCCCGCTGTTGCGCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
DB 916 ACAATACTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTATAATGAGAAAAATTAGTAG 975
QY 899 ACTTCGCGAGCTGAACAACCGCCAGCAGCTCTGGGAGGTGCAGCTGGGATCCCCC 958
DB 976 ATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACAC 1035
QY 959 ACCCGCGCGCCCTGAAGAAGAGAGCGCTGACCGTGTGCGAGCTGGGCGCAGCCCTACT 1018
DB 1036 ATCCCGCAGGCTTAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATT 1095
QY 1019 TCAGCGTGCCCTGGGACGAGGACTTCCGCAAGTACACCGCTTTCACCATCCCCAGCATCA 1078
DB 1096 TTTCACTTCCCTTAGATGAAGACTTCAGGAAGTATACTGCACTTTACCATACCTAGTATA 1155
QY 1079 ACAACGAGACCCCGCATCCGCTACAGTACACGCTGTGCCCCCAGGCTGGAAGGGCA 1138
DB 1156 ACAATGAGACACCGAGGATTAGATATCAGTACAAATGTGCTTCCACAGGGATGGAAGGAT 1215

QY 1139 GCCCGACGATCTTCCAGAGCAGATGACCAAGATCTCGAGCCCTTCCGGCCGCCCAACC 1198
DB 1216 CACCAGCAATATTCCTCAAGTAGCATGACAAAATCTTTAGAGCCTTTTAGAAAACAAATC 1275
QY 1199 CCAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGCGCAGGACCTGGAGATCG 1252
DB 1276 CAGACATAGTTATCTATCAATACATGATGATTTGTATGTAGATCTGACTTAGAATAG 1335
QY 1253 GCAGCAGCCGCGCAAGATCGAGAGCTGGCAAGCAGCCTGCTGCGCTGGGGCTTCACCA 1312
DB 1336 GCAGCATAGAACAAAATAGAGAGCTGAGACACATCTGTTGAGGTGGGGACTTACCA 1395
QY 1313 CCCCCGACGAAGACACCAAGAGAGCCGCCCTTCTGCCCCAT-----CGAGCTGCACC 1366
DB 1396 CACCAGACAAAACATCAGAAAGAACCTCCATTCTTTGGATGGTATGAACTCCATC 1455
QY 1367 CCAGCAAGTGGACCGTGCAGCCCATCGAGTGGCCGCGAGAGGAGCTGACCGCTGAACG 1426
DB 1456 CTGATAAATGGACAGTACAGCCTATAGTGTGCTGCCAGAAAAGACAGCTGGACTGTCAATG 1515
QY 1427 ACATCCAGAAAGTGGTGGGCAAGCTGAACCTGGCCGACGACATCTACCCGGCATCAAGG 1486
DB 1516 ACATACAGAAAGTGTGGGAAATTTGAATGGCAAGTCAATTTACCCAGGATTAAG 1575
QY 1487 TGGCCGACGTGTGCAAGCTGCTGCGGGGCCAAAGGCCCTGACCGACATCGTCCCTGTA 1546
DB 1576 TAAGGCAATTTATGTAACCTCTTAGAGGAACCAAGCACTTAACAGAAATACCACTAA 1635
QY 1547 CCAGAGGCGCGAGCTGGAGCTGGCGAGAACCCGCGAGATCTTGGCGGAGCCCGTGCAG 1606
DB 1636 CAGAAGAAGCAGAGCTAGACTGGCAGAAAACAGAGAGATTTCAAAGAACCACTAGTACATG 1695
QY 1607 GCCTGTACTACGACCCAGCAAGCAAGCTGCTGGCCGAGATCCAGACGACGAGGCCAGACC 1666
DB 1696 GAGTGTATTATGACCATCAAAAGACTTAATAGCAAAATACAGAGCAGGGGCAAGGCC 1755
QY 1667 AGTGGACCTACCAAGATCTACAGGAGCCCTTCAAGACCTTGAAGACCGGCAAGTACGCCA 1726
DB 1756 AATGGACATATCAAAATTTATCAAGAGCCATTTAAATCTGAAAACAGGAAATATGCAA 1815
QY 1727 AGATGGCACCAGCCACCAAGCAGCTGAGCAGCTGACCGAGGCGGTGCAGAAATCG 1786
DB 1816 GAATGAGGGTGGCCCACTAAATGATGTAAACAATTAACAGAGGAGTGCAGAAAATAA 1875
QY 1787 CCATGGAGCATCTGTGATCTGGGCAAGACCCCAAGTTCCGCCCTGCCATCCAGAAGG 1846
DB 1876 CCACAAAGCATAGTAATATGGGGAAGACTCTTAATTTAACTGCCCCATCAAAAGG 1935
QY 1847 AGACCTGGGAGACCTGGTGACCGACTACTTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1906
DB 1936 AAACATGGGAAACATGGTGACAGAGTATTTGGCAAGCCACTGATTCCTGTAGTGGAGT 1995
QY 1907 TCGTGACACCCCGCCCTGGTGAAGCTGTGGTACAGCTGGGAGGAGGCCCATCATCG 1966
DB 1996 TTGTTAATACCCCTCCCTTAGTGAATATATGTAACCATGTAACAGTGGAGTACCATAGT 2055
QY 1967 GCGCCGAGACCTTCTAGTGGCGCGCGCAACCCGAGACCAAGATCGCAAGGCCG 2026
DB 2056 GACAGAAACCTTCTATGTAGTGGGAGCTAACAGGAGACTTAATAGGAAAGCAG 2115
QY 2027 GCTACGTGACCGACCGCGCGCAGAAAGATCGTGAAGCTGACCGGAGACCAACCAACAGA 2086
DB 2116 GATATGTACTAATAGAGGAACAAAAGTTGTCACCTTAACCTGACACAAACAATCAGA 2175
QY 2087 AGACCGAGCTGACAGGCATCCAGCTGGCCCTGCGAGGACGCGGAGCGGAGGTGAACATCG 2146
DB 2176 AGACTGAGTTACAAAGCAATTTATCTAGCTTTGAGGATTCGGGATTAGAAGTAAACATAG 2235
QY 2147 TGACCGACAGCCAGTACGCCCTGGGATCATCCAGGCCAGCCGCAAGAGCGAGAGCG 2206
DB 2236 TAACAGACTACAAATATGCAATTAGGAATCAATCAAGCAACCAAGATCAAAAGTGAATCAG 2295
QY 2207 AGCTGTTGAACCAAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGG 2266

DB 2296 AGTTAGTCAATCAATAATAGAGCAGTTAATAAAAAAGGAAAGGTCTATCTGGCATGGG 2355
QY 2267 TGCCCCCCCACAAAGGCGATCGGGCGCAACGACGAGATCGACAGCTGCTGACGAGGGCA 2326
DB 2356 TACCAGCACACAAGGAATTTGGAGGAAATGAACAAGTAGATAAATTAGTCAGTCTGGAA 2415
QY 2327 TCCGCAAGGTGCTGTTCTCTCGGCGGCGCATCGAT 2358
DB 2416 TCAGGAAGTACTATTTTAGATGGAATAGAT 2447

RESULT 4
US-09-117-217-13
; Sequence 13, Application US/09117217
; Patent No. 6221578
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117,217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (750)..(2435)
; OTHER INFORMATION: Reverse Transcriptase
US-09-117-217-13

Query Match 44.0%; Score 1081.6; DB 4; Length 2601;
Best Local Similarity 67.9%; Pred. No. 4e-169;
Matches 1610; Conservative 0; Mismatches 734; Indels 28; Gaps 6;
QY 14 TGGCCGAGCGCATGAGCAGCGCCACCA---GCGCAACATCTCTGATCGCAGCGCAACT 70
DB 77 TGGCTGAAGCAATCAGCAAGTAACTACCAATTCAGCTACCAATATGATGAGAGGCAATT 136
QY 71 TCAAGGCCCCCAGCGCATCATCAAGTGTTCAACTCGGCAAGGAGGCGCACATCGCC 130
DB 137 TTAGGAACCAAGAAAGATTGTTAAGTGTTCATTTGTCGCAAGAGGGGCACACGCA 196
QY 131 GCAACTGCCCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGCGAAGGAGGGCCACCAGA 190
DB 197 GAAATTCAGGGCGCCCTAGGAAAAGGGCTGTGGAAATGTGGAAGAGGAGGACACAAA 256
QY 191 TGAAGGACTCCACGCGCGCCGCAAGGCAACTTCTTCCGAGGACCTGGCCCTTCCCCCAGG 250
DB 257 TGAAGATTGCTACTGAGAGACAGGCTAA-TTTTATAGGAGATCTGGCTTCTTACAAG 315
QY 251 GCAAGGCCCGGAGTTCGCCAGGAGAGCAACCGCGCAACAGCCCAACAGCCCGCGAGC 310
DB 316 GGAAGGCCAGGGAATTTCTTCAGAGCAGACCAAGCCACAGCCCAACAGAGAGAGC 375
QY 311 TGCAGGTGCCGG-----CGACAAACCCCGCAGCGAGGGCGCGCGCGCGCGAGGCA 364
DB 376 TTAGGTCTGGGTGAGAGCAACAACCTCCCGCTCAGAAAGCAGGAGCGGATAGACAAAGAA 435
QY 365 -----CCCTGAACCTCCCGCAGATCACCTGTGGCAGCGCCCTTGTGAGCATCAAG 418
DB 436 CTGTATCTTTAACTTCCTCAGGTCACTCTTTGGCAACGACCCCTCTGTACATAAGA 495
QY 419 TGGCGCGCCAGATCAAGGAGGCCCTTGTGGACACCGCGCGCGAGCACCCGTGTGAGG 478
DB 496 TAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 555


```
Db 3517 AATGGACATCAAAATTTATCAGAGCCATTTAAATCTGAACACAGGAAATATGCAA 3576
QY 1727 AGATGGCCACCCACACCAACAGCTGAAGCAGCTGACCGAGCGCGTGCAGAAGATCG 1786
Db 3577 GAATGAAGGTGCCCACACTAATGATGTGAACAATTAACAGAGGCGATACAAAAATAG 3636
QY 1787 CCATGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCCCGCTGCGCCATCCAGAAGG 1846
Db 3637 CCACGAAGAAGCATAGTAATATGGGGAAGAACTCTTAATTTAAATTACCCATCAAAAGG 3696
QY 1847 AGACCTGGGAGACTGCTGGACCGACTACTGGCAGCGCCACTCGATCCCGAGTGGAGT 1906
Db 3697 AATATGGGAAGCATGGTGACAGAGTATTTGGCAAGCCACTGGATTCTCTGAGTGGAGT 3756
QY 1907 TCGTGAACACCCCCCTCGTGGTGAAGCTGTGTACCAAGTGGAGAAGAGGCCCATCATCG 1966
Db 3757 TTGTCAATACCCCTCCCTTAGTGAAGTATGTATGTTACCAAGTATAGAGAAAGACATAG 3816
QY 1967 GCCTGAGAGCTTCTAGTGGAGCGCGCCCAAGAGATCGTGAGCCCTGACCGAGACCCACCAAGCGG 2026
Db 3817 GAGCAGAAATCTTCTATGTAGATGGGCGCAGCCAATAGGGAACCTAAATTAGGAAGACAG 3876
QY 2027 GCTACGTGACCGCGCGCGCAGAGATCGTGAGCCCTGACCGAGACCCACCAACAGCA 2086
Db 3877 GATATGTAAGTACAGAGGAGGACAAAAGTTGTCCCTTAACGGGACACAAACAAATCAGA 3936
QY 2087 AGACCGAGCTGACGGCATCCAGCTGGCCCTGCAGGACAGCGCGAGGTGAACATCG 2146
Db 3937 AGACTGAGTTACAAGCANNTCATCTAGCTTTGCAGATTCGGGATTAGAAGTAACATAG 3996
QY 2147 TGACGACAGCCAGTAGTACGCCCTGGGATCATCCAGGCCACCGCCGACAGAGAGGAGCG 2206
Db 3997 TGACAGACTCACAATATGCAATGGGAATCATCAAGCACAAACAGATAGAGTGAATCAG 4056
QY 2207 AGCTGTGACAGATCATCGACACTGATCAAGAGGAGGAGGTGTACCTGAGCTGGG 2266
Db 4057 AGTTAGTCACTCAATTAATAGACAGTTAATAAAAGGAAAAAGTCTACCTGGCATGGG 4116
QY 2267 TGCCCGCCCAAGAGGCATCGCGCGCAACAGCAGCATCGACAAGCTGGTGAGCAAGGGCA 2326
Db 4117 TACCAGCACAAAGAAATTTGGAGGAATGAACAAGTAGATGGTTGGTCAGTCTGGAA 4176
QY 2327 TCCGCAAGGTGCTGTTCTCGGAGCGGCATCGAT 2358
Db 4177 TCAGGAAAGTACTATTATTTAGATGGAATAGAT 4208
```

RESULT 6

```
US-08-188-583-5
; Sequence 5, Application US/08188583
; Patent No. 5851813
; GENERAL INFORMATION:
; APPLICANT: Desrosiers, Ronald C.
; TITLE OF INVENTION: PRIMATE LENTIVIRUS VACCINES
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,583
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 07/727,494
; FILING DATE: July 9, 1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/551,945
; FILING DATE: July 12, 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: Reg. No. 5851813 29,066
; REFERENCE/DOCKET NUMBER: 00246/079002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9709
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-188-583-5
```

```
Query Match 43.8%; Score 1075.2; DB 2; Length 9709;
Best Local Similarity 67.7%; Pred. No. 51e-168;
Matches 1606; Conservative 0; Mismatches 738; Indels 28; Gaps 6;

QY 14 TGCCGAGGCCATGAGCCAGGCCACCA---GCGCCCAACATCTCTGATGCGAGCGCAGCAACT 70
Db 1877 TGGCTGAAGCAATGAGCCAAGTAACAAATCCAGCTACCATAATGATACAGAAAGGCAATT 1936
QY 71 TCAAGGGCCCCAAGCGGCATCATCAAGTGTCTCACTTCGCGCAGAGGAGGCGGCACATCGCCC 130
Db 1937 TTAGGAACCAAGAAAGAAAGACTGTTAAGTGTTCATTTGTCGAAGAGAGGCGCATATAGCCA 1996
QY 131 GCAACTGCCCGCGCCCGCCCAAGAGGGCTGTGGAAGTCGCGAGGCGCGCGAGGAGGCGCACAGA 190
Db 1997 AATTTGCGAGGGCCCTTAGGAAAAGAGGCTGTTGGAATGTGGAAGAGAGGAGCACCAAA 2056
QY 191 TGAAGGACTGACCGAGCGCGCCCAACTTCTTCGCGAGGAGACCTTGGCCCTTCCCCCAGG 250
Db 2057 TGAAGATGTTACTGAGAGACAGGCTAA-TTTTATTAGGAAGATCTTGGCCTTCCCACAAG 2115
QY 251 GCAAGGCCCGGAGTTCGCCAGCGAGGAGAAACCGCGCCACACAGCCCGCCAGCGCGAGC 310
Db 2116 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGAGCAACACGCCCGCCACCAAGAGAGCG 2175
QY 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGGCGCGCGAGCGCGCGAGGCA 364
Db 2176 TTCAGGTTTGGGAAGAGAGACAACCTCCTCTCAGAGCAGGAGCGCATAGACAAGAA 2235
QY 365 -----CCCTGAACCTTCCCCCAGATCACCTGTGGCAGCGCCCGCTGTGTGAGCATCAAGG 418
Db 2236 CTGTATCCTTTAGCTTCCCTCAGATCACCTTTTGGCAGCGACCCCTCGTCAACAATAAGA 2295
QY 419 TGGCGCGCCAGATCAAGAGGCGCCCTGCTGGACACCGCGCGCGAGCAGACCGTGTGAGG 478
Db 2296 TAGGGGGCAATTTAAAGAGAGCTCTATTAGATACAGAGCAGATGATACAGTATTAGAAG 2355
QY 479 AGATGAGCCTGCCCGCAAGTGAAGCCCAAGATGATCGGCGCATCGCGCGCTTCATCA 538
Db 2356 AATGAATTTGCCAGGAAGATGGAAACCAAAATGATAGGGGAATTTGAGGTTTATCA 2415
QY 539 AGGTGCGCCAGTAGACGACCATCTCTGATCGAGATCTCTGCGCAAGAGGCCATCGGCACCG 598
Db 2416 AAGTAGACAGTATGATCAGATCTCATAGAAATCTCGCGACATATAAGCTATAGGTACAG 2475
QY 599 TGTGATCGGCCCCCGCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGTGGGCT 658
Db 2476 TATTAGTAGACCTACACCTGTCAACATAATTTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTTGAACTTCCCCCATCGAGCCCATCGAGACCGTGGCCGCGTGAAGCTCAAGCCCGCA 718
Db 2536 GCACCTTAAATTTCCCATTTAGTCCTATTGAGACTGTACCAGTAATAATTAAGCCAGGAA 2595
```

QY 719 TGGAGCGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAAATCAAGGCCCTGACCG 778
Db 2596 TGGATGCGCCCAAAATTAACAATGGCCATTTGACAGAGAAATAAAGCAATAGTAG 2655
QY 779 CCATCTCGGAGAGATGGAGAGGAGGCGAGATCAACCAAGATCGGCCCGGAGAACCCCT 838
Db 2656 AAATTTGTACAGAAATGGAAAAGGAGGAAAATTTTCAAAAATTTGGCCCTCAAAATCCAT 2715
QY 839 ACAACACCCCGTTCGCGCATCAAGAAAGAGGACACACCAAGTGGCGGAAGCTGGTG 898
Db 2716 ACAATACTCCAGTATTTGCCATAAAGAAAAGAGACAGTACTAAATGGAGAAAATAGTAG 2775
QY 899 ACTTCCCGAGCTGAACAAGCCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC 958
Db 2776 ATTTACAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTCAATTAGGAATACCAC 2835
QY 959 ACCCGCGCGGCTGAAGAAGAAGAGCGTGCCTGGAGCTGGGCGACGCCCTACT 1018
Db 2836 ATCCTGCAGGGTTAAAACACAGAAAATAACAGTAAACAGTACTGGATGTGGGGGATGCATATT 2895
QY 1019 TCAGCGTCCCGCTGGACGAGACTTCCGCAAGTACACCGCTTCCACCATCCCGACGATCA 1078
Db 2896 TTTAGTTCCTTTAGATAAAGACTTTCAGGAAGTATATCTGCAATTTACCATTCTAGTATAA 2955
QY 1079 ACAACGAGACCCCGGATCCGCTACCAAGTACACGCTGCTGCCCGAGGGCTGGAAGGGCA 1138
Db 2956 ACAATGAGACACCAAGGATTTAGATATATCAATGATGCTTCCACAGGGATGGAAGGAT 3015
QY 1139 GCCCGACGATTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCGCAACC 1198
Db 3016 CACGAGCAATATTTCCAGTGTAGCATGACAAAATACTTTAGAGCCCTTTAGAAAACAAAATC 3075
QY 1199 CCAGATCGTGTATACCA-----GGCCCCCTGTAGCTGGCAGCGACCTGGAGATCG 1252
Db 3076 CAGCATAGTCACTCTCAATACATGATGATTTGTATGAGGATGACTTGAATGAATAAG 3135
QY 1253 GCCAGCACCGCGCAAGATCGAGAGCTGGCGCAAGCACTCTGCTGGTGGGGCTTACCA 1312
Db 3136 GGCAGCATAGAACAAAATAAGAGAACTGACACAACTCTGTTGAGTGGGGATTACCA 3195
QY 1313 CCCCCACAGAACCCAGAGAGGCCCTTCTTCCCTGCCCCAT-----CGAGCTGACC 1366
Db 3196 CACGACACAAAACATTCAGAAAAGCTCCATTCCTTTTGGATGGGTATCAACTCCATC 3255
QY 1367 CGACAAGTGGACCTGCAGCCCATGAGCTGGCCGAGAGAGAGCTGGACCGTGAACG 1426
Db 3256 CTGATAATGGACAGTACAGCCCTATAGTCTGCCAGAAAAGGAGAGCTGGAGCTCAATG 3315
QY 1427 ACATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCCAGCCAGATCTACCCGGCATCAAG 1486
Db 3316 ACATACAGAAATTAGTGGGAAAATTTGAATTTGGCAAGTCAAGATTTATGCGAGGATTAAG 3375
QY 1487 TGGCCAGCTGTGCAAGCTCTGGCGGGCGCGAAGGCCCTGACCGACATCGTGGCCCTGA 1546
Db 3376 TAAGGCAATTTATGTAAACTTCTTAGGGGAACCAAGCACTAAACAGAAAGTAGTACCCTAA 3435
QY 1547 CCAGGAGGCGGAGCTGGAGCTGSCCGAGAACCGCGAGATCTTCCGCGAGCCCTGCAAG 1606
Db 3436 CAGAAGAGCAGAGCTAGAACTGCGAANAACCGGGAGATCTTAAAAGAACCGGTACATG 3495
QY 1607 GCGTGTACTACGACCCCGACGAGACCTGGTGGCGGAGATCCAGAGCAGGGCCGACGACC 1666
Db 3496 GAGTGTATTATGCCCATCAAGACATTAATAGCAGAAATACAGAGCAGGGCGCAAGGCC 3555
QY 1667 ATGGACCTACAGATCTACAGAGCCCTTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1726
Db 3556 AATGGACATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAA 3615
QY 1727 AGATGCCACCGCCACACACGACCTGAGCAGCTGACCGAGGCGCTGCAGAGATCG 1786
Db 3616 GAATGAAGGGTGGCCACACTTAATGATGTGAACAATTTAACAGAGGCGAGTACAAAAATAG 3675
QY 1787 CCATGGAGAGCATGCTGTGCGGCAAGACCCCAAGTTCCCGCCTCCCATCCAGAGG 1846

Db 3676 CCACAGAAAGCATAGTAATAATGGGAAAGACTCTTAATTTAAATTTACCCATACAAAGG 3735
QY 1847 AGACTCGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1906
Db 3736 AAACATGGGAAGCATGGTGGACAGAGTATTTGGCAAGCCACCTGGATTTCTGAGTGGAGT 3795
QY 1907 TCGTGAACACCCCGCTGGTGAAGCTGTGCTACCGAGCTGGAGAGAGAGCCCATCATCG 1966
Db 3796 TTGTCAATACCCCTTCCCTTAGTGAAGTTATGTTACCGATTAGAGAAAAGAACCCATAATAG 3855
QY 1967 GCGCGGAGACCTTCTAGCTGGACGCGCGCCACCGGAGACCAAGATCGCAAGGCCG 2026
Db 3856 GAGCAGAAACTTTCTATGTAGATGGGCGACCAATAGGGAACATAAATTAGSANAAGCAG 3915
QY 2027 GCTACGTGACCGACCGCGCGCGGCGGAGAGATCGTGTGAGCTGTACCGAGACCAACCCAGA 2086
Db 3916 GATATGTAACCTGACAGAGGAGAGAAAAGTTGTCCCGCTAACGCGACACAACAAATCAGA 3975
QY 2087 AGACCGAGCTGCAGCGCATCCAGCTGGCCCTGCAGGACGCGGAGGAGTGRACATCG 2146
Db 3976 AGACTGAGTTTACAAGCAATTCATCTAGCTTTGCAGGATTCGGGATTTAGAAATTAACATAG 4035
QY 2147 TGACCGCAGCAGCAGTACGCCCTGGGCGATCTCCAGCGCCGCGGACCAAGAGCGAGAGCG 2206
Db 4036 TGACAGACTCACAATATGCAATTTGGGAATCAITTCAGCACACACAGATGAAGTGAATCAG 4095
QY 2207 AGCTGGTGAACCAAGCATCATCGAGCAGCTGTATCAAGAGGAGAAAGTGTACTGAGCTGGG 2266
Db 4096 AGTTAGTCAGTCAAAATAATAGACAGTATAAATAAGGAAAAGTCTTACCTGGCATGGG 4155
QY 2267 TGCCCGCCCAAGGCGATCGCGGACGACGAGCATCGCAAGCTGCTGAGCAAGGCA 2326
Db 4156 TACCAGCACACAAAGGAATTTGGAGAAATGAACAAGTAGATGGGTTGGTCAGTCTGGAA 4215
QY 2327 TCCCAAGAGTCTGTTCTCTGGACGCGCATCGAT 2358
Db 4216 TCAGAAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 7
US-08-388-353-1
; Sequence 1, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Learmont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9703 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-388-353-1

Query Match 43.8%; Score 1075.2; DB 3; Length 9709;
Best Local Similarity 67.7%; Pred. No. 5.1e-168;
Matches 1606; Conservative 0; Mismatches 738; Indels 28; Gaps 6;

QY 14 TGCGCGAGGCATGAGCGGCCACCA---GCGCCACATCCTGTGTCAGCGCAGCAACT 70
DB 1877 TGCTGAAGCAATGAGCCAAAGTAACAAATCCAGCTACCATTAATGATACAGAAAGGCAATT 1936
QY 71 TCAAGGGCCCAAGCGGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGCCACATCGCCC 130
DB 1937 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATTTGTGCAAGAGGCCACATAGCCA 1996
QY 131 GCAACTGCCCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACACAGA 190
DB 1997 AAAATTTGACGGGCCCCCTAGGAAAAGGGCTGTTGGAATGTGAAAAGGAAGGACACCAA 2056
QY 191 TGAAGACTCACCGAGCGCCAGCCAACTTCTCCGCGAGGACCTGGCCCTTCCCCCAGG 250
DB 2057 TGAAGATTGCTAGAGACAGCGCTAA-TTTTATGGGAAGATCTGGCCCTTCCCACAG 2115
QY 251 GCAAGGCCCGAGTTCCTCCAGCGAGCAGACACCGCGCAACAGCCCAACAGCCCGGAGC 310
DB 2116 GGAAGGCCAGGAAATTTCTTCAGAGCAGACAGAGCCACAGCCCAACAGAGAGC 2175
QY 311 TGCAGTGCACGG-----CGACAACCCCGCAGCGAGGCGCGCGCGAGCGCCAGGGCA 364
DB 2176 TTCAGGTTTGGGAAGAGACAACTCCCTCTCAGAAGCAGGAGCGCATAGACAAGGAA 2235
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTGTGAGCATCAAGG 418
DB 2236 CTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAACAATAAGA 2295
QY 419 TGGCGCGCCAGATCAAGAGGCGCTGCTGGACACCGCGCGCCGACGACCGCTGCTGGAG 478
DB 2296 TAGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGATTAGAAG 2355
QY 479 AGATGAGCCTGCCCGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
DB 2356 AAATGAATTTGCCAGGAAGATGGAACCAAAAATGATAGGGGAAATTGGAGGTTTTATCA 2415
QY 539 AGGTGCGCCAGTACGACAGATCCTGTATCGAGATCTGCGCAAGAGGCCATCGGCACCG 598
DB 2416 AAGTAGACAGTATGATCAGATCACTCATAGAAATCTGCGGACATAAAGCTATAGGTACAG 2475
QY 599 TGCTGTGCGCCCGCCCGGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT 658
DB 2476 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGAACTTCCCGATCAGCCCATCGAGACCGTGGCCGCTGAGAGCTGAAGCCCGGCA 718
DB 2536 GCACCTTAAATTTTCCCATTTAGTCTATTGAGACTGTACAGTAAATTAAGCCAGGAA 2595
QY 719 TGCAGCGCCCAAGGTGAAGCTGCGCCCTGACCGGAGGAGAGATCAAGGCCCTGACCG 778
DB 2596 TGGATGCCCAAGAGTTAAACAATGGCCATTGACAGAAGAAAATAAAGCATTAGTAG 2655
QY 779 CCATCTGCGAGAGATGGAGAAGGGCAAGATCAACAGATCGGCCCGCGAGAACCCCT 838
DB 2656 AAATTTGTACAGAAATGGAAGAGGAGGAAAAATTTCAAAAAATTGGCCCTGAAATCCAT 2715
QY 839 ACAACACCCCGTGTTCGCCATCAAGAAAGGACAGCACCAAGTGGCGCGCAAGCTGGTG 898

DB 2716 ACAATACTCCAGTATTGGCCATAAAGAAAAAGACAGTACTAATATGAGAAAAATTAGTAG 2775
QY 899 ACTTTCGCGAGCTGAACAAGCAGCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCC 958
DB 2776 ATTTTCAGAGAACTTAATAGAGAACTCAAGATTTCTGGGAAGTTCAATTTAGGAATACCA 2835
QY 959 ACCCGCGCGCCTGAAGAAGAACGCGTGCCTGGAGCTGGGCGGAGCCCTACT 1018
DB 2836 ATCCTGCAGGTTAAACAGAAAAATCAGTAACAGTACTGGATGTGGCGCATGCAATT 2895
QY 1019 TCAGCTGCCCCCTGGAGGAGACTTCGCGAAGTACACCGCTTCACCATCCCCAGCATCA 1078
DB 2896 TTTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATACGTGATTTACCATACCTAGTATAA 2955
QY 1079 ACACAGAGACCCCGCATCCGTACCAAGTACACAGTGCCTGCCCGAGGGCTGGAAGGSCA 1138
DB 2956 ACAATGAGACACCGAGGATTAGATATCAGTACAATGTGCTTCCACAGGATGGAAGGAT 3015
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATTCCTGGAGCCCTTCCGGCGCCGCAACC 1198
DB 3016 CACCAGCAATATTCCAGTGTAGCATGACAAAAATCTTTAGAGCCTTTTAGAAAAACAATC 3075
QY 1199 CCGAGATCGTATCTACCA-----GGCCCCCTGTACGTGGCGAGCGACCTGGAGATCG 1252
DB 3076 CAGACATAGTCACTATCAATACATGATGATTGTTGTATGTAGGATCTGACTTAGAAATAG 3135
QY 1253 GCAGACACCGCGCAGATCGAGGAGCTGCGAAGCACTGCTGCGCTGGGGCTTCACCA 1312
DB 3136 GGAGCATAGAACAAAAATAGGGAAGTCTGAGCAACATCTGTTGAGGTGGGATTTACCA 3195
QY 1313 CCGCCGACAAAGACACCAAGAGCGCCCTTCTTCCCTCAT-----CGAGCTGCACC 1366
DB 3196 CACCAGACAAAAACATCAGAAAGAACTCCATTTCTTGTGATGGGTTATGAACCTCCATC 3255
QY 1367 CCGACAAAGTGGAGCTGCACCCCATCGAGCTGCCCGAAGAGGAGAGCTGGACCTGGAACG 1426
DB 3256 CTGATAAATGGACAGTACAGCCTATAGTGTGCCAGAAAAAGGACAGCTGGACTGTCAATG 3315
QY 1427 ACATTCAGAAAGCTGGTGGCAAGCTGAACTGGCGACAGATCTACCCCGGCATCAAGG 1486
DB 3316 ACATACAGAAATTTAGTGGGAAATTTGAATTTGGCAAGTCAAGTTTATGCGAGGATTAAG 3375
QY 1487 TGGCGCCAGCTGTCAAGCTGCTGCGGGGCCCAAGGCCCTGACCGACATCGTGCCCTGA 1546
DB 3376 TAAGGCAATTTATGTAACCTTCTTAGGGGAACCAAGCACTAACAGAAAGTAGTACCCTAA 3435
QY 1547 CCGAGAGCGCGAGCTGGAGCTGGCGGAGAACCCGAGATCTCTGCGGAGCCCGTGCACG 1606
DB 3436 CAGAAGAACGAGAGCTAGAACTGGCAGAAAACAGGGAGATTCTAAAAAGAACCGGTACATG 3495
QY 1607 GCGTGTACTACGACCCCGCAGCAAGGACCTGCTGGCGGAGATCCAGAACGAGGCCACGACC 1666
DB 3496 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAGCAGGGCAAGGCC 3555
QY 1667 AGTGGACCTACCAGATCTACCAGAGCCCTTCAAGAACCTCAAGACCCGCAAGTACGCCA 1726
DB 3556 AATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAAACAGGAAAAATATCAA 3615
QY 1727 AGATGGCACCCGCCACACACGACGTGAAGCAGCTGACCGAGGCGCGTCGAGAGATCG 1786
DB 3616 GAATGAAGGGTGGCCACACTAATGATGTGAACAATTTAACAGAGGCAAGTACAAAAATAG 3675
QY 1787 CCATGGAGAGCATCTGTGCTGGGCAAGACCCCAAGTTCCCGCTCCCATCCAGAAAG 1846
DB 3676 CCACAGAAACATAGTAATATATGGGAAAGACTCCTAAATTTAAATTTACCATACAAAGG 3735
QY 1847 AGACCTGGGAGACCTGCTGGACCGACTTACTGGCAGCGACCTGATCCCGCAGTGGAGT 1906
DB 3736 AAACATGGGAAGCATGCTGGACAGAGTATTGGCAAGCCACCTGGATTCTCTGAGTGGAGT 3795
QY 1907 TCGTGAACACCCCGCCCTGCTGAGCTGTGTTACCATCTGGAGAGGAGGCCCATCATCG 1966

Db 3796 TTGTCAATACCCCTCCCTTACTGAAGTTATGGTACCAGTTAGAGAAAGAACCCATAATAG 3855
QY 1967 GCGCCGAGACCTTCTAGCTGACGCGCGCCGCAACCCGCGAGACCAAGATCGGCAAGCCG 2026
Db 3856 GAGCAGAACTTTCTATGTAGATGGGAGCCCAATAGGGAACCTAAATTAGGAAAGCAG 3915
QY 2027 GCTACGTGACCGACCGGGCGGCGAGAGATCGTGAAGCTTACCCGAGACCAACCAACAGA 2086
Db 3916 GATATGTAACCTGACAGAGGAGACAAAAGTTGTCCCCCTAACCGACACACAAATCAGA 3975
QY 2087 AGACCGAGCTGCAGGCATCCAGCTGCGCCCTGCGAGGAGCGGCGAGGAGTGAACATCG 2146
Db 3976 AGACTGATTTACAAGCAATTCATCTAGCTTTTCAGGATTCGGGATTAGAAATTAACATAG 4035
QY 2147 TGACCCGACGACGCTACGCTCGGCTATCCAGGCGCCGAGGAGGAGGAGGAGGAGG 2206
Db 4036 TGACGAGCTCACATATGCTTGGATTCATTCAGGACACACCAAGATGAAGTGAATCAG 4095
QY 2207 AGCTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGGAGGAGGAGGAGGAGG 2266
Db 4096 AGTTAGTCAGTCAAAATATAGAGCAGTTAATAAAAAAGGAAAAAGTCTACCTGGCATGG 4155
QY 2267 TGCCCGCCCAAGGAGGATCGGCGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2326
Db 4156 TACAGCACACAAAGGAAATTTGGAGGAAATGAACAAGTAGATGGGTTGTCAGTCTGGAA 4215
QY 2327 TCCCAAGGTGCTTCTTCCGACGCGATCGAT 2358
Db 4216 TCAGGAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 8
US-08-488-551B-1

; Sequence 1, Application US/08488551B
; Patent No. 6015661
; GENERAL INFORMATION:
; APPLICANT: Nicholas J. Deacon
; APPLICANT: Dale A. McPhee
; APPLICANT: David Cooper
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 841
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,551B
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PM3864 (AU)
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: PM4002 (AU)
; FILING DATE: 21-FEB-1994
; APPLICATION NUMBER: PM0284 (AU)
; FILING DATE: 23-DEC-1994
; APPLICATION NUMBER: US 08/388,353
; FILING DATE: 14-FEB-1995
; APPLICATION NUMBER: PM3021/95
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FRANK S. DIGILIO
; REFERENCE/DOCKET NUMBER: 9606Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9709 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-488-551B-1

Query Match 43.8%; Score 1075.2; DB 3; Length 9709;
Best Local Similarity 67.7%; Pred. No. 5.1e-168;
Matches 1606; Conservative 0; Mismatches 738; Indels 28; Gaps 6;

QY 14 TGGCGGAGGCCATGAGCCAGGCCACCA---CGGCCAACATCCTGTGTCAGCGCAGCACT 70
Db 1877 TGGCTGAAGCAATGAGCAAGTAACAAATCAGCTACCATATATACATAAGAAAGCAATT 1936
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTTCAACTGGCGGAAGAGGGCCACATCGGCC 130
Db 1937 TTAGAACCAAGAAGAAAGACTGTTAAGTGTTCATTTGGCCAAAGAGGGCACATAGCCA 1996
QY 131 GCAACTGCGCGCCCCCGCGCAAGAGGCTGCTGGAAAGTGGCGCAAGAGGGCCACCAGA 190
Db 1997 AAAATTGCGAGGCCCTTAGGAAAAAGGCTGTGGAAATGTGAAAGGAAGGACACCAAA 2056
QY 191 TGAAGGACTGCACCGAGCGCCAGGCCCAACTTCTCCCGGAGGACCTGGCCTTCCCCAGG 250
Db 2057 TGAAGATTTGACTGAGAGACAGGCTAA-TTTTGTAGGGAAGATCTGGCCTTCCCAAG 2115
QY 251 GCAAGGGCCGCGAGTTCCCCAGCGAGAGAACCGCGCAACAGCCCCACACCGCCGAGC 310
Db 2116 GGAAGGCCAGGAATTTCTTCAGAGCAGACGAGCCACAGCCCCACACAGAGAGAGC 2175
QY 311 TGCAGTGGCGGG-----CGACAACCCCGAGCGCGCGCGCGCGCGCGCGCGCGCA 364
Db 2176 TTCAGTTTGGGGAAGAGACAACTCCCTCTCAGAAGCAGGAGCGCCATAGACAAGAA 2235
QY 365 -----CCCTGAACCTTCCCCAGATCACTTGGCAGCGAGCGCGCGCGCGCGCGCGCA 418
Db 2236 CTGTATCTTTAGCTTCCCTCAGATCACTTTGGCAGCGAGCGCGCGCGCGCGCGCGCA 2295
QY 419 TGGCGCGCAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 478
Db 2296 TAGGGGGCAATTAAGAGAGCTCTATTAGATACAGAGCAGATGATACATATTAGAG 2355
QY 479 AGATGAGCTGCCCGGCAAGTGGAAAGCCAAAGATGATCGGCGGATCGCGCGCTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGGAAACCAAAATGATAGGGGGAATTGGAGGTTTATCA 2415
QY 539 AGGTGCGCCAGTACGACGAGATCTGATCGAGATCTGCGGCAAGAGCCCATCGGCACCG 598
Db 2416 AAGTAGGACAGTATGATCAGATCTCATAGAAATCTGCGGACATTAAGCTATAGTACAG 2475
QY 599 TGTGATCGGCGCCCGCCCGTGAACATCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 658
Db 2476 TATTAGTAGGACCTACACCTGTCAACATATTGGAAGAAATCTGTTGACTCAGATGGCT 2535
QY 659 GCACCTGAACTTCCCATCAGCCCCCATCGAGACCGTGCCTGGAAGCTGAAGCGCGCGCA 718
Db 2536 GCACCTTAAATTTCCCATTAGTCTTATGAGACTGTACCAAGTAAATTTAAAGCCAGAA 2595
QY 719 TGGAGGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 2596 TGGATGCCCAAAAGTTAAACAAATGGCCATTGACAGAGGAAAAATAAAGCATTAGTAG 2655
QY 779 CCATCTCGGAGGAGTGAAGAGGAGGAGGAGATCAACCAAGATCGCGCGCGCGCGCGCGCT 838
Db 2656 AATTGTACAGAAATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2715
QY 839 ACAACACCCCGCTGTTCCGCTCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 898
Db 2716 ACAATACTCCAGTATTTGCCATTAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2775

Qy	899	ACTTCCGGAGCTGACACAGCCACCCAGGACTTCTGGAGGTCGACCTGGGCATCCCC	958
Db	2776	ATTTCAGAGAACTTAATAAGAACTCAAGATTTCTGGGAAGTTTCAATTAGGAATACCAC	2835
Qy	959	ACCCCGCCGCTGAAGAAGAAGAGCGTGACCGTCTGGACGTGGGCGACGCCCTACT	1018
Db	2836	ATCCTCGAGGTTTAAACAGAAAAATCAGTAACAGTACTTGGATGTGGCGATGCATATT	2895
Qy	1019	TCAGCTGTGCCCCGTGGAGAGACTTCCGGAAGTACACGCCCTTCCACATCCCAGCATCA	1078
Db	2896	TTTTCAGTCCCTTAGATAAAGACTTTCAGGAAGTATACTGCAATTTACCATACCTAGTATAA	2955
Qy	1079	ACAACGAGACCCCGCATCCGCTACCAGTACAAGCTGCTGCCCCAGGCTGGAAGGGCA	1138
Db	2956	ACAAATGAGACACACAGGATTAATATCAGTACATATGCTTCCACAGGATGAAAGGAT	3015
Qy	1139	GCOCAGCACTTTCCAGACGACATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACC	1198
Db	3016	CACGACGAATATTCCAGTGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAACAAATC	3075
Qy	1199	CCGAGATCTGTATCTACCA- ---- -GGCCCCCTGTAGCTGGCGAGCGACCTGGAGATCG	1252
Db	3076	CAGACATAGTCACTCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTTAGAATAG	3135
Qy	1253	GCCAGCACCGCCCAAGATCTGAGGAGCTGCGCAACACCTGCTGCGTGGGGTTCACCA	1312
Db	3136	GGCAGCATAGACAAAAATAGAGGAATCAGACAACATCTCTTGAGGTGGGATTTACCA	3195
Qy	1313	CCCCCGACAGACGACAGAGAGGCCCCCTTCTCTGCCAT-----CGAGCTGCACC	1366
Db	3196	CACGACAAAAAACATCAGAAGAACCCTCACTTCTTTGGATGGGTATTGAACTCCATC	3255
Qy	1367	CCGACAACTGGACCGTGCAGCCCATCGAGCTGCCCGAAGAGAGCTGGAACCGTGAACG	1426
Db	3256	CTGATAAATGGACAGTACAGCCTATAGTCTGCCAGAAAAGACAGCTGNACTGTCAATG	3315
Qy	1427	ACATCCAGAAGCTGTGGGCAAGCTGAACTGGCGCAGCCAGATCTACCCCGGCATCAAGG	1486
Db	3316	ACATACAGAAATTTAGTGGGAAAATTTGAATTGGCGAAGTCAGATTTATGCAAGGATTAAG	3375
Qy	1487	TGGCCAGCTGTCAAGCTGTCTGCGCGCGCCCAAGGCCCTGCACCGACATCGTGCCCTCGA	1546
Db	3376	TAAGGCAATTTATCTAACTCTTAGGGGAAACAAAGCACTAACAGAGTAGTACCACATA	3435
Qy	1547	CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTCGCGAGCCGCTGCAAG	1606
Db	3436	CAGAAGAAGCAGAGCTAGAACTGGCAGAAAAACAGGGAGATTTAAAAGAACCGGTACATG	3495
Qy	1607	CGTGTACTAGCACCCCGACGAGGACCTGGTGGCCGAGATCCAGAGACGAGGGCCACGACC	1666
Db	3496	GAGTGTATTATGACCCCATCAAAAGACTTAATAGCAGAAATACAGAACGAGGGCAGGGCC	3555
Qy	1667	AGTGGACCTACCAAGATCTACCAAGAGCCCTTCAAGAACCTGAAAGACGGCAAGTACGCCA	1726
Db	3556	AATGGACATATCAATTTATCAGAGCCATTTAAAACTGAAAACAGGAAATATGCCAA	3615
Qy	1727	AGATGCGCACCGCCCAACACAGAGCTGGAAGCTGACCGAGGCGGTGCAGAAGATCG	1786
Db	3616	GAATTGAAGGGTGGCCACACTAATGATGTGAAAACAAATTAACAGAGCGAGTACAAAAAATAG	3675
Qy	1787	CCATGGAGACATCTGTATCTGGGCGAGACCCCAAGTTCGCGCTGCCCATCCAGAAAGG	1846
Db	3676	CCACAGAAACGATAGTANTATGGGNAAGACTCTTAAATTTAAATTTACCATACAAAAGG	3735
Qy	1847	AGACTGGGAGACCTGGTGGACCGACTACTTGGCAGGCCACCTGGATCCCCGAGTGGGAGT	1906
Db	3736	AAACATGGGAAGCATGTTGCACAGAGTATTGGCAAGCCACCTGGATTCTGAGTGGGAGT	3795
Qy	1907	TCGTGAACACCCCCCTGGTGAAGCTGTGTACCAGCTGGAGAGAGAGGCCCATCATCG	1966
Db	3796	TTGTCAATACCCCTCCCTAGTGAAGTTATGTGTACCACTAGAGAAGAAAGAACCCATAATAG	3855
Qy	1967	GCGCCGAGACCTTTCTACGTGGAGCGCGCCGCAACCGCGGACCAAGATTCGGCAAGCGCG	2026

```

RESULT 9
US-09-309-572-15
; Sequence 15, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 15
; LENGTH: 9709
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-309-572-15

```

```

RESULT 9
US-09-309-572-15
; Sequence 15, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vector
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 9709
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type
US-09-309-572-15

```

Db 2176 TTCCAGGTTGGGGAAGAGACAACAACTCCCTCTCAGAGCAGGAGCGGATAGACAAGGA 2235
QY 365 -----CCCTGAACCTTCCCGCAGATCACCTGTGGGAGCGCCCGCTGGTGAGCATCAAGG 418
Db 2236 CTGTATCTTTCCTTCCCTCAGATCACTCTTTGGGAGGACCCCTCGTCACAAATAAGA 2295
QY 419 TGGGCGCCAGATCAAGGAGGCCCTCTGTGACACCGCGCCGACACCGCTGCTGGAG 478
Db 2296 TAGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGACGAGATGATACAGTATTAGAAG 2355
QY 479 AGATGAGCTTCCCGGCAAGTGAAGCCCAAGATGATCGGCGCATCGGCGGCTTCATCA 538
Db 2356 AATGAATTTGCCAGGAAGATGAACCAAAATGATAGGGGAATTGGAGGTTTATCA 2415
QY 539 AGTGGCCAGTACGACGAGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
Db 2416 AAGTAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2475
QY 599 TGCTGATCGGCCCGCCAGCCCGTGAACATCATCGGCGCAACATGCTGACCCAGCTGGCT 658
Db 2476 TATTAGTAGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGAATTTCCCATCAGCCCATCGAGACCGCTGCGGCTGAAGCTGAAGCCCGCA 718
Db 2536 GCACCTTAAATTTCCCATTTAGTCTTATTGAGACTGTACCCAGTAAATTAAGCCAGGA 2595
QY 719 TGAAGGCCCGCAAGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 2596 TGGATGCCCGCAAAAGTTAAACAATGGCCATTGACAGAGAAAATAAAGCATTTAGTAG 2655
QY 779 CCATCTCGAGGAGATGGAGAAGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT 838
Db 2656 AATTTGTACAGAAATGGAAAGAGAGAAATAATTCAAAAATTTGGCCCTGAAATCCAT 2715
QY 839 ACAACACCCCGCTTGGCCATCAAGAAGAGGACACCAAGTGGCGCAAGCTGGTG 898
Db 2716 ACAATCTCCAGTATTGGCCATAAGAAAGAAAGACAGTACTAAATGGAGAAATTAGTAG 2775
QY 899 ACTTCGCGAGCTGAACAAGCAGCCAGGACTTCTGGGAGGTGACGCTGGGATCCCG 958
Db 2776 ATTTACAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTCAATTTAGGAATACAC 2835
QY 959 ACCCGCGCGCTTGAAGAAGAGAGCTGACCGTGTGGAGCTGGGCGAGCGCTACT 1018
Db 2836 ATCTCGAGGTTAAACAGAAATAATCAGTACAGTACTGATGATGGGCGATGCAATTT 2895
QY 1019 TCAGCTGCGCCCTGGAGGAGCTTCGCAAGTACACCGCTTTCACCATCCCGCAGCATCA 1078
Db 2896 TTTCACTTCCCTTAGATAAAGACTTCAGGAAGTATACGCTATTTACCATACCTAGTATA 2955
QY 1079 ACACGAGACCCCGGATCCGCTACAGTACAACTGTGCGCCAGGCTGGAAGGGCA 1138
Db 2956 ACAATGAGACACCGGATTTAGATATCAGTACAACTGTGCTTCACAGGGATGGAAGGAT 3015
QY 1139 GCGCCAGCATCTTCAGAGCAGATGACCAAGATCTCTGGAGCCCTTCGCGCGCCGCAACC 1198
Db 3016 CACCGCAATATTCCAGTGTAGCATGCAAAAATCTTAGAGCCCTTTAGAAAACAATC 3075
QY 1199 CCGAGATCGTGTATACCA-----GGCCCCCTGTACGTGGGCGAGCAGCTGGAGATCG 1252
Db 3076 CAGACATAGTCATCATCAATACATGATGATTTGTATGATGATGATGATGATGATGATGAT 3135
QY 1253 GCAGACCGCGGCAAGATCGAGGAGTTCGCAAGCAGCTGCTGCGCTGGGCTTCACCA 1312
Db 3136 GGCAGCATAGCAAAAATAGGGAATGAGCAACATCTGTTGAGGTGGGATTTTACCA 3195
QY 1313 CCGCCGACAGAGACACCAAGGAGCCCGCTTCCTGCGCCAT-----CGAGCTGCACC 1366
Db 3196 CACGACAAAAAACATCAGAAAGAACTCCATTTCCCTTTGATGGGTATTGAATCCATC 3255
QY 1367 CCGCAAGTGGCGCTGCACCCATCGAGCTGCGCGGAGAGGAGCTGGACCGCTGAACG 1426

Db 3256 CTGATAAATGGACAGTACAGCCTATAGCTGTCGAGAAAAAGACAGCTGGACTGTCAATG 3315
QY 1427 ACATCAGAAAGCTGGTGGGCAAGCTGAACCTGGGCCACCCAGATCTACCCCGGATCAAGG 1486
Db 3316 ACATACAGAAATTAGTGGGAAATTTGAATTTGGGCAAGTCAGATTTATGCAAGGATTAAG 3375
QY 1487 TGGCGCAGCTGTGCAAGCTGTGCGCGCCCAAGGCCCTGACCGACATGCTGCCCTGA 1546
Db 3376 TAAGGCAATTTATGAACTTCTTAGGGAACCAAGCACTAACAGAGATAGTACCATAA 3435
QY 1547 CCAGGAGCGGAGCTGGAGCTGGCGAGAAACCGCGAGATCTCTGCGAGCCCGCTGCACG 1606
Db 3436 CAGAAGAGCAGAGCTAGAAGCTGGCAGAAAACAGGGAGATTCTAAAAAGAACCGGTACATG 3495
QY 1607 GCGTGTACTACGACCCAGCAAGGACCTGTGGCCGAGATCCAGAACGAGGCCACGACC 1666
Db 3496 GAGTGTATTATGACCCCATCAAAAGACTTAATAGCAGAAATACAGAGCAGGGGCAAGGCC 3555
QY 1667 AGTGGACCTACCGATCTACCGAGGCGCTTCAAGAACTTCAAGACCGGCAAGTACGCCA 1726
Db 3556 AATGGACATATCAATTTATCAAGAGCCATTAAATAATCTGAAGACAGGAAATATGCAA 3615
QY 1727 AGATGCCACCGCCCAACACGACGCTGAAGCAGCTGACCGAGGCGCTGCAGAAAGATCG 1786
Db 3616 GAATGAAGGGTGGCCACACATTAATGTGAACAATTAACAGAGGCAATACAAAAATAG 3675
QY 1787 CCATGGAGACATCGTCTGCGGCAAGACCCCAAGTTCCGCGCTGCCATCCAGAGG 1846
Db 3676 CCACAGAAAGCATAGTAAATATGGGGAAGACTCCTAAATTTAAATTTACCCATACAAAGG 3735
QY 1847 AGACTGGGAGACCTGTGTGACCGACTACTTGGCAGGCGACCTGTGATCCCGAGTGGGAGT 1906
Db 3736 AATCATGGAGACATGTGTGACAGAGTATTGGCAAGCCACCTGTGATTCCTGAGTGGAGT 3795
QY 1907 TCGTGAACACCCCGCTGCTGTAAGCTGTGTACAGCTGGAGAGAGGCCATCATCG 1966
Db 3796 TTGTCATACCCCTCCCTTAGTGAAGTTATGGTACCAGTTAGAGAAAGAACCCCATATAG 3855
QY 1967 GCGCGGAGACCTTCTAGCTGGACGCGCGCCCAACCGGAGACCAAGATCGGAAGGCCG 2026
Db 3856 GAGCAGAAACTTTCTATGTAGATGGGCGAGCAATAGGGAACATAAATAGGAAGAGCAG 3915
QY 2027 GCTACGTGACCGCGCGGCGGAGAGATGCTGAGCCTGACCGGAGACCAACCAACAGA 2086
Db 3916 GATATGTAATGACAGAGGAGACAAAAGTTGTCCTTAACCGACACACAAATCAGA 3975
QY 2087 AGACCGAGCTGACGCGCATCCAGCTGGCCCTGCGAGGACAGCGGAGGTGAACATCG 2146
Db 3976 AGACTGATTTACAAGCAATTCATCTAGCTTTGAGGATTCGGGATTTAGAACTAAACATAG 4035
QY 2147 TGACCGACGCGCAGTACGCGCTGGGATCATCCAGGCCCGCCGACAGAGCGAGCGG 2206
Db 4036 TGACAGACTCACAATATGCTATTTGGGAATTCATTAAGCACAACACAGATGAAGTGAATCAG 4095
QY 2207 AGCTGGTGAACCAAGATCATCGAGCAGCTGATCAAGAGAGGAGAGGTGTACCTGAGCTGG 2266
Db 4096 AGTTAGTCAGTCAATTAATAGAGCAGTTAATAAAGGAAAGATCTACCTGGCATGGG 4155
QY 2267 TGCCCGCCCAAGGGGATCGGCGGCAAGCAGAGATGCAAGCTGGTGAAGAGGCA 2326
Db 4156 TACCAGCACAAAGAAATTTGGAGGAAATGAACAGTAGATGGGTTGGTCACTGCTGGA 4215
QY 2327 TCCGCAAGGTGCTTCTCTGACCGCATCGAT 2358
Db 4216 TCAGGAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 10

US-08-935-312-13
; Sequence 13, Application US/08935312
; Patent No. 6207455
; GENERAL INFORMATION:
; APPLICANT: CHANG, Lung-Ji

ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 015280-249000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 15581 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: - 15581
LOCATION: 1..15581
OTHER INFORMATION: /note= "pNLnSg11"

US-08-646-538-35

Query Match 43.8%; Score 1075.2; DB 3; Length 15581;
Best Local Similarity 67.7%; Pred. No. 5.3e-168;
Matches 1606; Conservative 0; Mismatches 738; Indels 28; Gaps 6;

QY 14 TGCCGAGGCGCATGAGCCAGGCGACCA---GCCGCCAACATCCTGTGATGCGAGCGCAGCAACT 70
DB 1877 TGGCTGAAGCAATGAGCCAGTAACAAATCCAGCTACCAATAATGATACAGAAAGGCAATT 1936
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTTCACACTGCGCGCAAGGAGGCGGCACATCGCCC 130
DB 1937 TTAGGAACCAAGAAAGACTGTAAAGTGTTCAAATTGTGCAAGAGAGGCGCATAGGCCA 1996
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGCAGGCGCAAGGAGGCGCACAGA 190
DB 1997 AAAATTGACGGGCCCTAGGAAAAGGCTGTGGAAATGTGGAAGAGGAGACACAAA 2056
QY 191 TGAAGACTGCAACGAGCGCCAGGCCAACCTTCTCCGCGAGGACCTGGCCCTCCCGCCAGG 250
DB 2057 TGAAGATTGTACTGAGACAGAGCTAA-TTTTTTAGGGAAGATCTGGCCCTCCACAAAG 2115
QY 251 GCAAGGCCCGGAGTTCCCGAGCGAGCAGACAGACCGCGCCACACGCCCCACCGCGGAGC 310
DB 2116 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGAGCCACAGACGCCCCACAGAGAGAGC 2175
QY 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGGCGCGCGCGAGCGCGAGGCA 364
DB 2176 TTCAGGTTGGGGAAGAGACAACTCCCTCTCAGAGCAGGAGCGCGATAGACAGGAA 2235
QY 365 -----CCCTGAATTCGCCCGAGATCAACCTGTGGCAGCGCCCGCTGGTGAGCATCAAGG 418
DB 2236 CTGTATCTTTAGCTTCCCTCAGATCACCTCTTTGGCAGCGACCCCTCGTCACAATAAAGA 2295
QY 419 TGGCGCGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGAGCACCGCTGCTGGAGG 478
DB 2296 TAGGGGGGCAATTAAGAGAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 2355
QY 479 AGATGAGCCTGCGCGGCAAGTGAAGGCCAAGATGATGCGCGCATCGCGCGCTTCATCA 538
DB 2356 AATGAAATTGCGCAGNAGATGGAACCAAAATGATAGGGGAATTGGAGGTTTATCA 2415
QY 539 AGGTGCGCCAGTACGACCCAGATCCTGATCGAGATCTGCGGCAAGAGGCGCATCGGCACCG 598
DB 2416 AAGTAGACAGTATGATCAGATCTCATAGAAATCTCGCGACATAAAGCTATAGGTACAG 2475
QY 599 TGTGATCGGCCCGCCCGTGAACATCATCGCGCGCACATCTGTACCCAGCTGGGCT 658
DB 2476 TATTAGTAGGACCTTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGTAATTTCCCATCAGCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCGGCA 718
DB 2536 GCACTTTAAATTTTCCCATAGTCTTATGAGACTGTACCAGTAAATTTAAGCCAGGAA 2595
QY 719 TGGACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGGAGGAGAAATCAAGGCCCTGACCG 778

DB 2596 TGGATGCCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAATAAAGCATTTAGTAG 2655
QY 779 CCATCTCGGAGGAGATGGAGAAGGCGAAGATCCCAAGATCGGCCCGGAGAACCCCT 838
DB 2656 AAATTTGTACAGAAATGGAAAAGAGGAAAAATTTCAAATAATGGGCGCTGGAATCCAT 2715
QY 839 ACAACACCCCGCTGTTGCGCATCAAGAAGAGGACACACAGTGTGCGCAAGCTGGTGG 898
DB 2716 ACAATACTCCAGTATTTGCCATAAAGAAAAAGACACTACTAAATGAGAAAAATTAGTAG 2775
QY 899 ACTTCGCGAGCTGAACAAGCGCCAGGACTTCTGGGAGGTGCGACCTGGGCGATCCCC 958
DB 2776 ATTTTCAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGATACAC 2835
QY 959 ACCCGCGCGGCTGAAGAAGAGAGCGTGCAGCGTGTGGAGCTGGGCGGCGACCCCTACT 1018
DB 2836 ATCTCGAGGGTTAAACACAGAAAAATCAGTAACAGTACTGGGATGTGGGCGATGCATATT 2895
QY 1019 TCAGCGTGCCTCGGAGGAGACTTCCGCAAGTACACCGCTTCACCATCCCCAGCATCA 1078
DB 2896 TTTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATACCTGATTTACCATACCTAGTATA 2955
QY 1079 ACAACGAGACCCCGGATCCGCTACCACTACACGTGTGCTGCCCGCAGGCTGGGAAGGCA 1138
DB 2956 ACAATGAGACACCCAGGATTAGATATCAGTACAATGTGCTTCCACAGGATGGAAGGAT 3015
QY 1139 GCCCGAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCGCAACC 1198
DB 3016 CACGAGCAATATTTCCAGTGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAACAAATC 3075
QY 1199 CCGAGATCGTGATCTACCA-----GGCCCCCTGTGACGTGGGCGGAGCAGCTGAGATCG 1252
DB 3076 CAGACATAGTCACTATCAATACATGATGATTTGTATGTAGGATCTGACTTACAAATAG 3135
QY 1253 GCAGACCCGCGGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGTGGGGCTTCACCA 1312
DB 3136 GGCAGCATAGAAAAAATAGAGGAAGTACAGAACATCTGTTGAGGTGGGGATTTACCA 3195
QY 1313 CCCCCGACAGAGACACCAAGAGGAGCCCGCTTCCCTGCCCAT-----CGAGCTGCACC 1366
DB 3196 CACGACAAAAAACAATCAGAAAGAACCTCCATTCTTTGGATGGGTATGAACTCCATC 3255
QY 1367 CCAGCAAGTGCAGCTGCAGCCCATCGAGCTGCCGAGAGAGAGAGCTGGACCGTGAACG 1426
DB 3256 CTGATAATGGACAGTACAGCTTATAGTGTGCCAGAAAGGACAGCTGACTGCTCAATG 3315
QY 1427 ACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCGCAGCCAGATCTACCCGGCATCAAGG 1486
DB 3316 ACATACAGAAATAGTGGGAAAAATGAATTGGGCAAGTCAAGATTTATGCGAGGATTAAG 3375
QY 1487 TGGCGCAGCTGTCAAGCTGCTGCCGCGCGCCAGGCGCTGACCGAGATCGTGGCCCTGA 1546
DB 3376 TAAGGCAATTTATGTTAAACTTCTTAGGGGAACCAAGACACTAACAGAGTAGTACCACTAA 3435
QY 1547 CCAGAGGAGCGCGAGCTGGAGCTGGCCGAGAACCGGAGATCTCGCGGAGCGCCGTGACG 1606
DB 3436 CAGNAGAGCAGAGTAGAAGTGCAGAAACACAGGAGATCTTAAAGAACCCGTACATG 3495
QY 1607 GCGTGTACTACGACCCCGCAAGGAGCTGTGTGCGCGAGATCCAGAAAGCGGCGCACGACC 1666
DB 3496 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAAGCAGGCGCAAGGCC 3555
QY 1667 AGTGGACCTACCAGATCTACCAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1726
DB 3556 AATGGACATATCAAAATTTATCAGAGCCATTTAAAAATCTGANAACAGGAAAAATATGCAA 3615
QY 1727 AGATGCGACCGCCACACCAACGAGCTGAAGCAGCTGACCGGCGCGTGCAGAAAGTACG 1786
DB 3616 GAATGAAGGTGCCCCACACTAATGATGTGAACAAATTAACAGAGGCGAGTACAAAAATAG 3675
QY 1787 CCATGGAGAGCATGTGATCTGGGCGAAGACCCCAAGTTCCCGCTGCCCATCCAGAGG 1846

Db 3676 CCACAGAAAGCATAGTAATATGGGAAAGAGCTCTAAATTTAAATTAACCATACAAAGG 3735
QY 1847 AGACTGGAGAGCTGGTGACACCACTACTGGCAGGCCACTGGATCCCGAGTGGAGT 1906
Db 3736 AAACATGGGAAGCATGGTGACAGAGATTGGGAAGCCACTGGATTCTGAGTGGAGT 3795
QY 1907 TCCTGAACACCCCTGGTGAAGCTGGTACCAAGCTGGAGAGGAGCCCATCATCG 1966
Db 3796 TTGTCAATACCCCTCCCTAGTGAAGTATGTTACCAAGTATGAGAAAGACCCATAATAG 3855
QY 1967 GCGCCGAGACCTTCTAGTGGAGCGGCCGCCCAACCCGAGACCAAGATCGGAAGCCG 2026
Db 3856 GAGCAGAAACTTTCTATGTATGATGGGACGCCAATAGGGAACCTAAATTAGGAAGACG 3915
QY 2027 GCTACGTGACCGCGCGCGGCGCAGAGATCGTGAAGCTTGACCGGAGACCAACACAGA 2086
Db 3916 GATATGTAACGTACAGAGGAGGACAAAGATTGTCCCTTAACGGACACAAACATCAGA 3975
QY 2087 AGACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACAGCGGAGGTGAACATCG 2146
Db 3976 AGACTGATTAACAAGCAATTCATCTAGCTTGCAGGATTCCGGATTAGAAGTAACATAG 4035
QY 2147 TGACCGACACCGAGTAGCGCCCTGGGCATCATCCAGGCCCGCCGACGAGAGCG 2206
Db 4036 TGACAGACTCACAAATATGCAATGGGAATCATTCGAAGCACAACAGATAAGAGTGAATCAG 4095
QY 2207 AGCTGGTGAACCATCATCGACGAGCTGATCAAGAGGAGAGAGGTGTACTGAGCTGGG 2266
Db 4096 AGTTAGTCACTCAANTATAGACGAGTTAATAAAGGAAAGTCTACTGGCATGGG 4155
QY 2267 TGCCCGCCCAACAGGCGATCGCGGCCAACGACGAGATCGACAAGCTGGTGAGCAAGGCA 2326
Db 4156 TACCAGCACAAAGGAATTTGGAGGAATCAACAAGTAGATGGTGGTGGCTGGAA 4215
QY 2327 TCAGCAAGTGTCTTCTCGGAGCGGATCGAT 2358
Db 4216 TCAGGAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 13

US-09-503-222-35
; Sequence 35, Application US/09503222
; Patent No. 6265548
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Gaitanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,222
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,538
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677

; REFERENCE/DOCKET NUMBER: 015280-249000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: - 1.15581
; LOCATION: 1..15581
; OTHER INFORMATION: /note= "pNlnsg11"
US-09-503-222-35

Query Match 43.8%; Score 1075.2; DB 4; Length 15581;
Best Local Similarity 67.7%; Pred. No. 5 3e-168;
Matches 1606; Conservative 0; Mismatches 738; Indels 28; Gaps 6;
QY 14 TGGCCGAGCCATGAGCCAGGCCACCA---GCGCCACATCTCTGATCGAGCGCAGCAACT 70
Db 1877 TGGCTGAAGCAATGAGCAAGTAACAAATCCAGCTACCATAATGATACAGAAAGGCAATT 1936
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTCGGCAAGGAGGGCCACATCGCCC 130
Db 1937 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATTTGTGCAAGAGAGGCGCATATGCCA 1996
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAAAGGCTGCTGGAAGTGGCGAAGGAGGGCCACCAGA 190
Db 1997 AAATTCAGGGCCCTTAGGAAAGAGGCTGTTCGAATGTGGAAGAGAGGACACCCAA 2056
QY 191 TGAAGGACTCGACGAGCGCCAGCCCAACTTCTTCCGAGGAGCTTGGCTTCCCCCAGG 250
Db 2057 TGAAGATTTACTTGAGAGACAGCTAA-TTTTATTAGGAAGATCTGGCTTCCCAACAAG 2115
QY 251 GCAAGGCCCGCGAGTTCCCGCAGGAGCAAGCCGCAACAGCCGCAAGCGCCGAGCGCGGCA 310
Db 2116 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGAGCCAAAGCCCGCCAGCAAGAGAGC 2175
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGAGCGCGCGAGGCA 364
Db 2176 TTCAGGTTTGGGAAGAGAGACAACACTCCCTCTCAGAAGCAGGAGCGCGATAGACAAGAA 2235
QY 365 -----CCCTGAACCTTCCCGCCAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAG 418
Db 2236 CTGTATCTCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAGA 2295
QY 419 TGGCGCGCCAGATCAAGAGGCGCTGCTGGACACCGCGCGCGAGCAGCGCGCGCGCGCGAGG 478
Db 2296 TAGGGGGCAATTAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGATTAGAG 2355
QY 479 AGATGAGCTTCCCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCAATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGAAACCAAAATATAGGGGAAATTGGAGGTTTTTATCA 2415
QY 539 AGGTGCGCCAGTACGACCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
Db 2416 AGTAGGACAGTATGATCATAGATACTCATAGAAATCTCGGACACATAAAGCTATAGGTACAG 2475
QY 599 TGTGTATCGGCGCCCGCCCGTGAACATCATCGGCGCAACATGCTGACCCAGCTGGGCT 658
Db 2476 TATTAGTAGGACCTACACCTGCTCAACATAATTGGAAGAAATCTGTGTGACTCAGATTGGCT 2535
QY 659 GCACCTGAACCTTCCCGCATCAGCCCGCATCGAGACCGTGGCCGTGAAGCTGAAGCCCGCA 718
Db 2536 GCACCTTAAATTTTCCCATTTAGTCTTATTGAGACTGTACCAAGTAAATTAAGCCAGAA 2595
QY 719 TGGACGGCCCGCAAGGTGAAGCAGTGGCCCTTGACCGGAGAGAGATCAAGGCCCTGACCG 778
Db 2596 TGGATGGCCCAAGGTTAAACAATGGCCCTTGACAGAAGAAAAAATAAAGCATTTAGTAG 2655

Db	1944	TTAGGAACCAAGAAAGACGTGTTAAAGTGTTTCAATTGGCGCAAGAAGGGCACAATAGCCA	2001
Qy	131	GCAACTGCCCGCCCCCGCCCAAGAAGGGCTGCTGGAAGTGGCGCAAGAGAGGGCCACAGA	190
Db	2004	AAATTGCAGGGCCCCTAGGAAAAAGGGCTGTTGAGATGTGGAAGGAAGGACACCAA	2063
Qy	191	TGAAGACTGCACCGAGCGCGAGGCCAAGTCTTCCGCGAGGAGACTGGCCTTCCCCGAGG	250
Db	2064	TGAAAGATTGCATGTAGACACAGGCTAA-TTTTTTTAGGGAAGATCTGGCCTCTCTACAAG	2122
Qy	251	GCAAGGCCCGCGAGTTTCCCGACGAGCAGAACCGCGCCCAACAGCCCCACACAGCGCGGAGC	310
Db	2123	GGAGGCCAGGGATTTTCTTCAGAGCAGACCAAGCCACAGCCCCACAGAAAGAGAGC	2182
Qy	311	TGCAGGTGCGCGGCG-----ACAACCCCCGCGAGCGAGGGCGCGCGCGCGCCAGGGCA	364
Db	2183	TTCAGGTTTGGGAGGAGAAAAACAATCCCTCTCAGAAGCAGGAGCGGATAGACAAGAA	2242
Qy	365	-----CCCTGAACTTCCCCAGATCACCTGTGCGACGCGCCCCCTGGTGAGGATCAAGG	418
Db	2243	CTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAAGACCCCTCGTCCAAATAGGA	2302
Qy	419	TGGCGGCCAGATCAAGAGAGGCCCTGCTGGACACCGCGCGCGACGACACCTGCTGGAGG	478
Db	2303	TAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG	2362
Qy	479	AGATGAGCCTGCCCGCAGTGGAGGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCA	538
Db	2363	AAATGAAATTTGCCAGGAAATGAAACCAAAATGATAGGGGAATTTGGAGGTTTATCA	2422
Qy	539	AGGTGCCCACTACGACCAGATCCTGATCGAGATCTGCGGCAAGAAGGCCATCGGCACCG	598
Db	2423	AAGTAAGACAGTACGATCAGATACCTGTAGAAATCTGTGGACATAAAGCTATAGGTACAG	2482
Qy	599	TGCTGATGCGCCCCCCCCCGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGCT	658
Db	2483	TATTAGTAGACCTACACCTGTCAACATAAATTTGGAAGAAATCTGTTGACTCAGATGGTT	2542
Qy	659	GCACCCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCTGAAAGTGAAGCCCGGCA	718
Db	2543	GTACTTTAAATTTCCCACTTAGTCTATTGAAACTGTACCAAGTAAATTTAAAGCCAGGAA	2602
Qy	719	TGGAGGCCCCAGGTGAAGCAGTGGCCCTTGACCGAGGAGAAAGATCAAGGCCCTGACCG	778
Db	2603	TGGATGCCCCANAAGTTAAGCAATGGCCATTGACAGAAGAAAAATTTAAAGCAATTAGTAG	2662
Qy	779	CCATCTCGAGGAGATGGGAAGAGGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT	838
Db	2663	AGATATGTACAGAAATTTGAAAGGAGGAAATTTCAAANATTTGGGCCTGAAATTCAT	2722
Qy	839	ACAACACCCCGTGTTCGCCATCAGAAGAAGAGCAGCACCAAGTGGCGCAAGCTGGTGG	898
Db	2723	ACAATACTCCAGTATTGTGTATAAGAAAAAGACAGTACTAAATTTGGAGAAAACTAGTAG	2782
Qy	899	ACTTCCCGAGCTGAACAGCGCACCCAGGACTCTGGGAGTGCAGCTGGGATCCCCC	958
Db	2783	ATTTACAGAACTTATAAAGAACTCAAGACTTCTGGGAAGTTTCAGTTAGGAATACCAC	2842
Qy	959	ACCCCGCGGCCTGAAGAAGAAAGAGCGTGACCGTCTGGAGCTGGGCGACGCTACT	1018
Db	2843	ACCCCGCAGGTTTAAAAAGAAAAAATCAGTAACAGTATTGGATGTGGTGATGCATACT	2902
Qy	1019	TCAGCGTGGCCCTTGGAGGAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGATCA	1078
Db	2903	TTTCAGTTCCTTTAGATAAAGACTTTAGAAAGTATTACTGCAATTTACCATACTAGTATAA	2962
Qy	1079	ACAACGAGACCCCGCATCCGCTACAGTACAACGCTGCTGCCCCAGGGCTTGAAGGGCA	1138
Db	2963	CAATAGACACACAGGGATTAGNATCAGTACATATGCTGCCACAGGGATTGAAGGAT	3022
Qy	1139	GCCCCAGCATTTCCAGAGCAGCATGACCAAGATCCTTGGAGCCCTTCCGGCCCGCAACC	1198
Db	3023	CACCAGCAATATTTCCAAGTGTAGCTGACAAAAATCTTTAGAGCTTTTAAAAACAGAACT	3082

QY	1199	CCGAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGGCACGACCTCGGAGATCG	1252
DB	3083	CAGACATAGTTATCTATCAATACATGGATGATTTGTGTGTAGGATCTGACTTAGAANTAG	3142
QY	1253	GCCAGCACCGCCAGATCGAGGAGCTGGCAGACACCTGCTGCCTGGGCTTGCCACCA	1312
DB	3143	GGCAGCATAGAACAAAATAGAGGAATGAGACAGACTCTGTTAGGTGGGATTTACCA	3202
QY	1313	CCCCGACAAGAAGCACCAAGAGGCCCTTCCTGCCCAT-----CGAGCTGCACC	1366
DB	3203	CACCAGACAAAACATCAGAAAGACCTCCATTCTTTGGATGGGTTATGAACCTCCATC	3262
QY	1367	CGACAAAGTGGACCGGTGCAGCCCATCGAGTGGCCGAGAAAGAGAGCTGGACGCTGAACG	1426
DB	3263	CTGATAAATGGACGTPACAGCCTTATAATGCTGCCAGAAAAGACAGCTGGACTGTCAATG	3322
QY	1427	ACATCCAGAGCTGGTGGCGCAAGCTCAACTGGSCCAGCCAGATCTACCCGGCATCAAGG	1486
DB	3323	ACATACAGAAGTTAGTGGGAAATTAATTTGGCAGAGTCAGATTTATGCAAGGATTAAAG	3382
QY	1487	TGCGCCAGCTGTGCAAGCTCTGCGCGGGGCCAAGCCCTGACCGACATCGTGCCCTGTA	1546
DB	3383	TAAAGCAGTTATGTAACCTCTTAGAGGAACCAAGCACTAACAGAAGTAATACCACTAA	3442
QY	1547	CGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGGCGGACCCCGTGCACG	1606
DB	3443	CAGAAGAAGCAGAGCTAGAACTGGCAAGAACAGGAGATTTCTAAAAGAACCACTACATG	3502
QY	1607	GGGTGTACTACAGACCCACGCAAGGACCTGGTGCCGAGATCCAGAAGCGGGCCACGACC	1666
DB	3503	AGATATATTATGACCCATCAAAAGACTTAGTAGCAGAAATACAGACGAGGCGCAGGCC	3562
QY	1667	AGTGGACCTACCAAGTCTACAGGAGCCCTTCAAGAACCTGAAGACGGCAAGTACGCCA	1726
DB	3563	AATGGACATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAACAGGAAAGTATGCAA	3622
QY	1727	AGATGGCACCGCCACACCAAGCAGCTGAGCAGAGCTGACCGAGCCGTCGCAAGAGTCG	1786
DB	3623	GGATGAGGGGTGCCACACTAATGATGTAAACAGTTTACAGAGGCAAGTGCAAAAAGTAT	3682
QY	1787	CCATGGAGAGACTCGTGATCTGGGGCAAGACCCCAAGTTCCGCGCTGCCCATCCAGAAGG	1846
DB	3683	CCACAGAAAGCATAGTAATGGGAAAGATTCTTAAATTTAACTTACCCTACAAAAAGG	3742
QY	1847	AGACTGGAGAGCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT	1906
DB	3743	AAACATGGGAAGCATGGTGGATGGAGTATTGGCAAGCTACTGGATTTCTTGAGTGGGAGT	3802
QY	1907	TCGTGAACACCCCCCTGGTGAAGCTGTGTTACAGCTGGAGAAGGAGCCCATATCG	1966
DB	3803	TGTGTAATACCCCTCCCTTAGTGAATATATGTTACCAGTTAGAGAAGAACCCTAGTAG	3862
QY	1967	GGGCCGAGACTTCTACGTGGAGCGGCCGCCCAACCGCGAGACCAAGTCGGCAAGGCCG	2026
DB	3863	GAGCAGAACTTCTATGTAGATGGGCAGCTAATAGGGAGACTAAATAGGAAAAACAG	3922
QY	2027	GCTACTGACCGACCGGGCCGCGCAGAAGATCTGTAGCCCTGACCGAGACCAACCAACAGA	2086
DB	3923	GATATGTTACTGACAGAGAAAGCAAAAAGTTGTCTCCATAGCTGACACAAACAAATCAGA	3982
QY	2087	AGACCAGCTGCAGGCCATCCAGCTGSCCTGCAGACACGCGCAGGAGGTGAACATCG	2146
DB	3983	AGACTGAAATTACAGCAATTCATCTAGTTTGAGGATTCGGGATTTAGAAGTAAACATAG	4042
QY	2147	TGACCACACGCAAGTACGCCCTGGGCCTCATCCAGGCCACCGCCGACAGAGCGAGACGG	2206
DB	4043	TAAACACTCACAATATGCATTAGGAATCATTTCAAGCACAACCAAGATAAGTGAATCAG	4102
QY	2207	AGCTGGTGAACCAAGATCATCGAGCAGCTGATCAAGAAAGGAAAGGTACTCTGAGCTGGG	2266
DB	4103	AGTTAGTCTAGCTCAAAATAATAGACAGTTAATAAAAAGGAAAGAGTCTACTCGCATGGG	4162

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:45:48 ; Search time 107.404 Seconds
(without alignments)
10843.369 Million cell updates/sec

Title: US-09-610-313-32

Perfect score: 2457

Sequence: 1 gtcgacgcaccatggccga.....gggctagcaccgggaattc 2457

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1607.4	65.4	4307	10	US-09-999-183-2
2	1539.2	62.6	8366	10	US-09-872-733-6
3	1503.8	61.2	4338	10	US-09-872-733-1
4	1137	46.3	2467	10	US-09-872-733-3
5	1081.6	44.0	2601	10	US-09-735-487-7
6	1081.6	44.0	2601	10	US-09-735-487-9
7	1081.6	44.0	2601	10	US-09-735-487-11
8	1081.6	44.0	2601	10	US-09-735-487-13
9	1081.6	44.0	4307	10	US-09-999-183-1
10	1081.6	44.0	9719	10	US-09-737-190A-1
11	1029.2	41.9	8933	10	US-09-943-286-3
12	1029.2	41.9	8933	10	US-09-943-286-4
13	1029.2	41.9	8933	10	US-09-943-286-9
14	1022.8	41.6	9544	10	US-09-798-675-4
15	1022.8	41.6	9544	10	US-09-798-675-5
16	818.4	33.3	2348	9	US-10-097-997-1
17	805.8	32.8	9793	9	US-09-886-156-56
18	805.8	32.8	9793	9	US-09-886-150-56
19	805.8	32.8	9793	9	US-09-886-149-56

20	805.8	32.8	9793	9	US-09-886-159-56	Sequence 56, Appl
21	699.8	28.5	2507	10	US-09-872-733-2	Sequence 2, Appl
22	657.2	26.7	12379	10	US-09-991-258-14	Sequence 14, Appl
23	654	26.6	1323	10	US-09-991-258-15	Sequence 15, Appl
24	425.4	17.3	948	10	US-09-756-551A-5	Sequence 5, Appl
25	243.8	9.9	1800	9	US-10-003-035-58	Sequence 58, Appl
26	239.2	9.7	4516	10	US-09-968-355-24	Sequence 24, Appl
27	216.4	8.8	12523	10	US-09-991-258-1	Sequence 1, Appl
28	215.6	8.8	1476	10	US-09-991-258-4	Sequence 4, Appl
29	172.6	7.0	507	9	US-10-003-035-56	Sequence 56, Appl
30	163.2	6.6	1482	10	US-09-818-443-4	Sequence 4, Appl
31	163.2	6.6	1532	10	US-09-818-443-1	Sequence 1, Appl
32	157.8	6.4	1503	10	US-09-968-355-25	Sequence 25, Appl
33	156.6	6.4	1752	10	US-09-968-355-16	Sequence 16, Appl
34	156.6	6.4	4594	10	US-09-968-355-15	Sequence 15, Appl
35	138.2	5.6	9464	10	US-09-738-847-1	Sequence 1, Appl
36	133.6	5.4	1496	9	US-10-003-035-17	Sequence 17, Appl
37	130.6	5.3	13049	9	US-09-996-073-1	Sequence 1, Appl
38	130.6	5.3	13049	10	US-09-231-235-1	Sequence 1, Appl
39	130.6	5.3	13049	10	US-09-797-518A-1	Sequence 1, Appl
40	112.6	4.6	2310	10	US-09-476-242-21	Sequence 21, Appl
41	112.6	4.6	2535	10	US-09-476-242-13	Sequence 13, Appl
42	111	4.5	2328	10	US-09-476-242-6	Sequence 6, Appl
43	111	4.5	2523	10	US-09-476-242-15	Sequence 15, Appl
44	108.6	4.4	2517	10	US-09-476-242-16	Sequence 16, Appl
45	108.6	4.4	2517	10	US-09-476-242-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-999-183-2
; Sequence 2, Application US/09999183
; Patent No. US20020147169A1
; GENERAL INFORMATION:
; APPLICANT: MITROPHANOUS, et al
; TITLE OF INVENTION: In Vivo Selection Method
; FILE REFERENCE: 674523-2009
; CURRENT APPLICATION NUMBER: US/09/999,183
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: PCT/GB00/02136
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 9912965.2
; PRIOR FILING DATE: 1999-06-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: SeqWin99
; SEQ ID NO 2
; LENGTH: 4307
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Codon optimised gagpol sequence
US-09-999-183-2

Query Match	65.4%	Score 1607.4;	DB 10;	Length 4307;
Best Local Similarity	81.7%	Pred. No. 2.4e-256;		
Matches 1939;	Conservative	0;	Mismatches 406;	Indels 28; Gaps 6;
QY	12	CATGCCGAGGCCATGAGCCAGG---CCACCAGCGCCCAACCTCTGATGCAGCGCAGCAA	68	
Db	1086	CCTGCTGTGAGCCATGAGCCAGGTGACCACTCCGCTACCATCATGATGCAGCGCGCAA	1145	
QY	69	CTTCAAGGGCCCAAGCGCATCATCAAGTGTCAACTGCGGCAAGGAGGCCACATGCG	128	
Db	1146	CTTTCGGAACCAACCAAGCAAGTCGTCAAGTGTCAACTGTGCAAGAAAGGACACAGC	1205	
QY	129	CGCAACTGCGCGCGCCCGCCCGGCTGCTGGAAGTGCAGGAGGAGGCGCCACCA	188	
Db	1206	CGCAACTGCGAGGCGCCCTAGGAAAGAGGGTGTGGAATTCGGCAAGAGGACCA	1265	
QY	189	GATGAAGGACTGCACCGAGCGCGCAGGCCAACCTCTTCGCGGAGGACCTGGCTTCC	248	

Db 1266 GATGAAGACTGTACTGAGAGACAGGCTAA-TTTTTAGGGAAGATCTGGCCCTTCCTACA 1324
QY 249 GGGCAAGGCCGCGAGTTCCCAAGCAGAGCAAGAACCGCGCAACAGCCGCCACCAAGCCGCGA 308
Db 1325 AGGAAGGCCAGGGAATTTCTTCAGAGCAGACCAAGAGCCCAACAGCCCAACAGAGAGA 1384
QY 309 GCTGAGGTGCGGG-----CGACAAACCCCGCAGAGCGCGCGCGAGGCCAGGG 362
Db 1385 GCTTCAGGTCTGGGTAGAGACAACAACCTCCCTCAGAAGCAGGAGCCGATAGACAAGG 1444
QY 363 CA-----CCCTCAACTTCCCCCAGATCACCTTGGCAGCGCCCTCGTGGAGCATCAA 416
Db 1445 AACTGATCTTTAACTTCCTCAGATCACTTTGGCAACGACCCCTCGTCAATAAA 1504
QY 417 GGTGGCGGCAGATCAAGAGGCCCTGTGGACACCGCGCGCAGACACACCGTGTGGA 476
Db 1505 GATAGGGGCGACCTCAAGAGGCTCTCTGTGGACACCGGAGCAGACACCGTGTGGA 1564
QY 477 GGAGATGAGCCTGCCGGCAAGTGGAAAGCCCAAGATGATCGGCGGATCGGCGCTTCAT 536
Db 1565 GGAGATGTCTTGTCCAGGGCGCTGGAAAGCCGAAGATGATCGGGGAATCGCGGTTTCAT 1624
QY 537 CAAGGTGCGCAGTACGACAGATCCTGATCGAGATCTCGGCAAGAGCCATCGGCAC 596
Db 1625 CAAGGTGCGCAGTATGACAGATCTCTATCGAAATCTGGCGCCCAAGGCTATCGGTAC 1684
QY 597 CGTGCTGATCGGCCCAACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGG 656
Db 1685 CGTGCTGGTGGCCCCACACCCGCTCAACATCATCGGACGCAACTGTTGACGCAGATCGG 1744
QY 657 CTCACACCCCGTGTGCCATACGCCCATCGAGACCGTGCCTGAAAGTGAAGCCCGG 716
Db 1745 TTCCAGCGCTGAACCTTCCCATTAAGCCCTATCGAGACGGTACCGGTGAAAGTGAAGCCCGG 1804
QY 717 CATGGAGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAAGATCAAGCCCTGAC 776
Db 1805 GATGGAGGCCCGGAAGCTCAAGCAATGGCCATTGACAGAGGAAGATCAAGCACTGGT 1864
QY 777 CGCATCTGCGAGGATGAGAAAGAGGCAAGATCACCAAGATCGGCCCGCAGAACCC 836
Db 1865 GGAGATTTGCACAGATGGAAGAAGGAGGAATCTCCAAGTTGGGCTGAGAACCC 1924
QY 837 CTACACACCCCGTGTGCCATCAAGAAAGAGACAGACCAAGTGGCGCAAGCTGGT 896
Db 1925 GTACAAACACCGCGGTGTGCCAATCAAGAAGAGGACTCGACAAATGGCGCAAGCTGT 1984
QY 897 GGACTTCCGGAGCTGNACAAAGCCACCCAGACTTCTGGGAGGTGCAGCTGGGCATCC 956
Db 1985 GGACTTCCGGAGCTGAACAAGCGCACGAAGACTTCTGGGAGGTTCAGCTGGGCATCC 2044
QY 957 CCACCCCGCGGCTGAAGAAGAAGAGCGTGAACGCTGCGAGCTGGCGGACGCCTA 1016
Db 2045 GCACCCCGAGGCTGAGNAGAAAGAAATCCGTGACCGTACTGGATGGGTGATGCCTA 2104
QY 1017 CTTTACGCTGCCCTGGACGAGACTTCGCGAAGTACACCGCTTCACCAATCCCGCAGCAT 1076
Db 2105 CTTCTCCGTTCCCTGGACGAACCTTCAGGAAGTACACTGCCTTCACAATCCCTTCGAT 2164
QY 1077 CAACAAGGAGACCCCGGATCCGCTACCAATGATACAGTGTCTGCCCGAGGGCTGGAAGG 1136
Db 2165 CAACAAGGAGACCCGGGATTCGATATCAGTACAAAGCTGCTGCCCGCAGGGCTGGAAAGG 2224
QY 1137 CAGCCCGCAGCATCTTCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCGCGCCGCGCAA 1196
Db 2225 CTCCTCCGCAATCTTCAGAGTAGCATGACCAAAATCTCTGGAGCCCTTCGCGAAACAGAA 2284
QY 1197 CCCCAGATCGTGTATACCA-----GGCCCCCTGTACGTGGGAGCGACCTTGGAGAT 1250
Db 2285 CCCCAGATCGTGTATATCAGTACATGATGATGACTTGTACGTGGGCTGCTATCTAGAT 2344
QY 1251 CGGCCAGCACCCGCGCAAGATCGAGAGCTGCGCAGACCTGCTGGCTGGGCTTCAC 1310
Db 2345 AGGGCAGCACCCGCAAGATCGAGAGCTGGCCAGACCTTGTGAGGTGGGACTGAC 2404

QY 1311 CACCCCGCACAAAGCACACAGAGGAGCCCTTCTTCCCTCCCAT-----CGAGCTGCA 1364
Db 2405 CACACCGCACAAAGCACACAGAGGAGCCCTTCTTCCCTCTGGATGGTTACGAGCTGCA 2464
QY 1365 CCCCAGCAAGTGAACCGTGCAGCCCATCGAGCTGCCCGAAGAGAGAGCTGGACCGTGAA 1424
Db 2465 CCGTGCAAAATGGACCGTGCAGCCTATCTGCTGCCAGAGAAAGACAGCTGGAGTGCAA 2524
QY 1425 CGACATCCAGAAGCTGGTGGCAAGCTGAACTGGCCAGCCAGATCTACCCCGGCATCAA 1484
Db 2525 CGACATACAGAAGCTGGTGGGAAGTTGAACTGGCCAGTFCAGATTTACCCAGGATTA 2584
QY 1485 GGTGCGCCAGCTGTGAAGCTGTGCGCGCGCAAGCCCTGACCCACATCTGCTGCCCT 1544
Db 2585 GGTGAGCAGCTGTGCAAACTCTCCGCGGAACCAAGGCACTCACAGAGTGTATCCCT 2644
QY 1545 GACCGAGGAGCGCAGCTGAGCTGGCCGAGAACCGCGAGATCTTGGCGAGGCCGTGCA 1604
Db 2645 AACCGAGGAGCGCAGCTCGAATCTGCGAGAAACCGAGAGATCTTAAAGGAGCCGCTGCA 2704
QY 1605 CGCGGTGTACTAGACCCCAAGCAGGACCTGGTGGCCGAGATCCAGAAGCAGGCGCCACA 1664
Db 2705 CGCGGTGTACTATGACCCCTCCAAAGGACCTGTATCGCCGAGATCCAGAAGCAGGCGCAAGG 2764
QY 1665 CCAGTGGACCTTACAGATCTTACAGGAGCCCTTCAAGAACCTTGAAGACCGCAAGTACGC 1724
Db 2765 CCAGTGGACCTATCAGATTTACAGGAGCCCTTCAAGAACCTTGAAGACCGCAAGTACGC 2824
QY 1725 CAAGATGCGCACCGCCCAACCAACGACAGCTGAAGCAGCTGACCGAGCCCTGCAAGAAT 1784
Db 2825 CCGGATGAGGGGTGCCACACTAACGACGTCAAGCAGCTGACCGAGCCCTGCAAGAAT 2884
QY 1785 CGCATGGAGAGCATGTGATCTGGGCAAGACCCCAAGTTCGGCTGCCATCCAGAA 1844
Db 2885 CACCCCGAAGGATCGTGATCTGGGAAAGACTTCTTAAGTTCAAGTGCCTCCATCCAGAA 2944
QY 1845 GGAGACTGGGACACCTGTGTGGACCACTACTGCGAGCCACCTGGATCCCGAGTGGGA 1904
Db 2945 GGAACCTGGGAACCTGTGTGGACAGATATTGGCAGCCACCTGGATCTCTGAGTGGGA 3004
QY 1905 GTTTCGTGAACACACCCCGCTGTGTGAAGCTGTGTACAGCTGGAGAAAGAGCCCATCAT 1964
Db 3005 GTTTCGTGAACACCCCGCTGTGTGAAGCTGTGTACAGCTGGAGAAAGAGCCCATAGT 3064
QY 1965 CGCGCGGAGACCTTCTACCTGGACGCGCCGCAACCGCGAGACCAAGATTCGCAAGGC 2024
Db 3065 GCGCGCGAAGACCTTCTACGTGGATGGGCGCGCTAACAGGGAGACTAAGCTGGGCAAGC 3124
QY 2025 CGCTACGTGACCGACCGCGCGCGAGAGATCGTGAGCTGACCGAGACCAACCAACCA 2084
Db 3125 CGATACGTCACTAACCGGGCAGACAGAGTTGTACCCCTCACTGACACCAACCAACCA 3184
QY 2085 GAAGACCGAGCTCAGGCCATCCAGCTGGCCCTGCAAGCAGCGGAGAGGAGTGAACAT 2144
Db 3185 GAAGACTGAGCTCAGGCCATTTACCTCGCTTGTGACAGACTTCGGGCTGGAGGTGAACAT 3244
QY 2145 CGTGACCGACAGCCAGTACGCCCTGGCATCATCCAGGCCCGCCCGCAGAGCGAGAG 2204
Db 3245 CGTGACAGACTCTCAGTATGCCCTGGCATCATTCGAAGCCCGCAGCAGAGTGAATC 3304
QY 2205 CGAGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAAAGGTGTACCTGAGCTC 2264
Db 3305 CGAGCTGGTCAATCAGATCATCGAGCAGCTGATCAAGAAGAAAGGTCTATCTGGCGCTG 3364
QY 2265 GGTGCCCGGCCCAAGGGCATCGCGCGCAACGACGAGATCGACAAGCTGTGTGAGCAAGG 2324
Db 3365 GGTACCCGCCCAAAAGGCATTTGGCGCAATGAGCAGGTGCAACAAGCTGTCTCGGCTGG 3424
QY 2325 CATCCGCAAGGTGCTTCTTCTGGACGGCATCGA 2357
Db 3425 CATCAGGAAGGTGCTATCTCTGTGATGGCATCGA 3457

RESULT 2

US-09-872-733-6
; Sequence 6, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of the construct pCMVgagpolBtkan containing a CMV
; OTHER INFORMATION: promoter, a HIV gag/pol gene and a kanamycin
; OTHER INFORMATION: resistance gene
US-09-872-733-6

Query Match 62.6%; Score 1539.2; DB 10; Length 8366;
Best Local Similarity 79.9%; Pred. No. 4.1e-245;
Matches 1896; Conservative 0; Mismatches 448; Indels 28; Gaps 6;

QY 14 TGGCGGAGGCATGAGCAGCCACAGC---GCCACATCTCTGATGCGCGGAGCAACT 70
Db 1857 TGGCGGAGGCATGAGCAGCGTGCAGCACTCGCGGACCATTAATGTCAGAGGCAACT 1916
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTCTTCAACTCGCGCAAGGAGGGCCACATCGCCC 130
Db 1917 TCCGGAACACAGCGAAGATCGTCAAGTCTCAATTGTGCAAGAGGCAACCGCCA 1976
QY 131 GCACTGCGCGCGCCCGCGCAAGAGGGTCTGGAAGTGCAGAGGCGGCAACAGCA 190
Db 1977 GGAATGCGCGGCGCCCGGCAAGAGGGTCTTGGAAATGTGAAAGGAGGACACAAA 2036
QY 191 TGAAGGACTGCAAGCGGCGAGCCCAACTTCTTCGCGAGGACCTGGCCCTTCCCGCAGG 250
Db 2037 TGAAGATTGTCTGAGAGACAGCTAA-TTTTGTAGGAAGATCTGGCCCTTCTTCAAG 2095
QY 251 GCAAGGCGCGGAGTTCCTCCAGCGAGCAGAACCGCGCCCAACAGCCCGCAGCGCGGAGC 310
Db 2096 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGAGCCACAGCCCGCAGAGAGAGC 2155
QY 311 TGCAGTTCGCGG-----CGACNAACCGCGAGCGGCGCGCGCGCGCGCGCGGCA 364
Db 2156 TTCAGGCTCGGGGTAGAGAACAACTTCCCGCTCAGAAAGCAGGCGGATAGACAAGAA 2215
QY 365 -----CCCTGAATCTCCCGAGATACCTGTGGCAGCGCCCGCTGGTGAGCATCAAG 418
Db 2216 CTGTATCTTTAATCTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAGTAAGGA 2275
QY 419 TGGCGGCGCAGATCAAGGAGCGCTGTGGACACCGCGCGCGCGCGCGCGCGCGGAGG 478
Db 2276 TCGGGGGCAACTCAAGGAAGCGCTGCTCATACAGGAGCAGATGATACAGTATTAGAG 2335
QY 479 AGATGAGCTGCGCGCAAGTGGAGCCCAAGATGATCGCGCGCATCGCGGGCTTCATCA 538
Db 2336 AAATGAGTTTGGCAGGAAGATGGAACCAAAAATGATAGGGGGGATCGGGGGCTTCATCA 2395
QY 539 AGGTGCGCAGTAGCAGCATCTGATCAGATCTCGCGCAAGAGGCGCATCGGCACCG 598
Db 2396 AGGTGAGCAGTAGCAGCATCTGATGAAATCTGTGACATAAAGCTATTAGGTACAG 2455

QY 599 TGCTGATGGCCCGCCACCGCTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT 658
Db 2456 TATTAGTAGACCTACACCTGTACATTAATTGGAGAATAATCTGTTGACCCAGATCGGCT 2515
QY 659 GCACCTGAACTTCCCATCAGCCCCATCGAGACCGTGCCTGTAAGCTGAAGCCCGGCA 718
Db 2516 GCACCTTGAATTTCCCATCAGCCCCATTATGAGACGCTGCCCTGAAGTTGAAGCCGGGA 2575
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTTGACCG 778
Db 2576 TGGACGCGCCCAAGGTGAAGCAATGGCCATTGACGAAAGAGAGATCAAGGCCCTTAGTCG 2635
QY 779 CATCTCGCAGGAGATGGAGAAGGAGGAGGCAAGATCACCAAGATCGCCCGCGGAAACCCCT 838
Db 2636 AATCTGTACAGAGATGGAGAGGAGGAGAGATCAGCAAGATCGGGCTTGAGAACCCCT 2695
QY 839 ACAACACCCCGCTGTTCGCCATCAAGAAGAAGAGCAGCACCAGTGGCGCAAGCTGGTGG 898
Db 2696 ACAACACTCCAGTCTTCGCAATCAAGAAGAAGAGAGTACCAAGTGGAGAAAGCTGGTGG 2755
QY 899 ACTTCGCGGAGCTGAACAAGCGCACCCAGGACTTCTGGAGAGTGCAGCTGGGCATCCCC 958
Db 2756 ACTTCAGAGAGCTGAACAAGAGAACTCAGGACTTCTGGGAAGTTTCTGAGGATCCCCAC 2815
QY 959 ACCCGCGCGCTGAAGAAGAAGAGAGCGTGCACCTGTCTGACGCTGGGCGACGCTACT 1018
Db 2816 ATCCCGCTGGTGTGAAGAAGAGAGTCAAGTGCAGTGTGGTGTGAGTGTGCTACT 2875
QY 1019 TCAGCTGCCCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACATCCCCAGCATCA 1078
Db 2876 TCTCCGCTTCCCTTGGACGAGGACTTCAGGAAGTACACTGCCCTTACGATACCTAGCATCA 2935
QY 1079 ACAACGAGACCCCGCGCATCCGCTACCAGTACAAGTGTGCTGCCCGCAGGCTGGAAGGCA 1138
Db 2936 ACAACGAGACCCCGCGCATCCGCTACCAGTACAAGTGTGCTGCCACAGGATGGAAGGAT 2995
QY 1139 GCGCCAGACTTTCAGAGCAGCAGTACCAAGATCTGGAGCCCTTCGCGCGCCGCAACC 1198
Db 2996 CACGAGCATCTTCAAGCAGCATGACCAAGATCCTGGAGCCCTTCGCAAGCAAAACC 3055
QY 1199 CGGAGATCTGTATCTACCA-----GGCCCCCTGTACTGGGCGAGGACCTGGAGATCG 1252
Db 3056 CAGACATCTGTATCTATCATAGTACATGGAGGACCTCTACCTAGGAAGTACCTGGAGATCG 3115
QY 1253 GCAGCAGCCGCGCAAGATCGAGGAGCTCGCAAGCAGCTGCTGCGCTGGGCTTCACCA 1312
Db 3116 GCGAGCAGAGCAAGATCGAGGAGCTGAGACAGCATCTGTTGAGTGGGAGCTGACCA 3175
QY 1313 CCGCCGACAAAGAACCCAGAGGAGGCGCCCTTCTGCCCCAT-----CGAGCTGCACC 1366
Db 3176 CACGAGCAAGAAGCAGCAGGAGGAACTCCCTTCTGTTGGATGGCTTACGAACCTGCATC 3235
QY 1367 CGGCAAGTGGACGTGCGAGCCCATCGAGCTGCCGAGAGGAGAGCTGGACCTGGAAGC 1426
Db 3236 CTGACAAGTGGACGTGCGAGCCCATCGTGTCTGCTGAGAAGGACAGCTGGAGCTGTAAGC 3295
QY 1427 ACATCCAGAAGCTGTGGGCAAGCTGAACCTGGCGCAGCAGCATCTACCCCGCATCAAG 1486
Db 3296 ACATCAGAAGCTGTGGGCAAGTTGAACCTGGGCAAGCAGATCTACCCAGCATCAAG 3355
QY 1487 TGGCGCAGCTGTGCAAGCTGTCTGCGCGGCGCAAGGCCCTGACCGACATCTGCGCCCTGA 1546
Db 3356 TTAGCAGCTGTGCAAGCTGTCTGCGAGGAAACCAAGGCACTGACAGAGTGTATCCACTGA 3415
QY 1547 CGGAGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGCGGAGCCCGCTGACAG 1606
Db 3416 CAGAGGAAGCAGAGCTGAGAACTGGCAGAGAACCGGAGATCTCTGAAGGAGCAGTACATG 3475
QY 1607 GCGTGTACTACACCCCGCAGCAGGAGCTGGTGGCGGAGATCAGAAAGCAGGCGCCAGACC 1666
Db 3476 GAGTGTACTACACCCCGCAGCAGGAGCTGTATCGCAGAGATCAGAAAGCAGGCGCAAGGCC 3535
QY 1667 AGTGGACCTTACCAGATCTTACCAGGAGGCCCTTCAAGAACCTTGAAGACCGCGCAAGTACGCCA 1726

Db 3536 AATGGACCTACCAAACTTACCAGAGCCCTCAAGAACCTGAAGACAGGCAAGTACGCCAA 3595
QY 1727 AGATGCCACCCGCCACCAACGACCTGAAGCAGCTGACCGAGGCGGTGACGAAGATCG 1786
Db 3596 GGATAGGGGTGCCACCAACACATGTGAAGCAGCTGACAGAGGCGATGACGAAGATCA 3655
QY 1787 CCATGGAGAGCATCGTGATCTGGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAGG 1846
Db 3656 CCACAGAGACATCGTGATCTGGGGCAAGACTCCCAAGTTCAAGCTGCCCATACAGAAGG 3715
QY 1847 AGACTGGGAGACCTGGTGACCGACTACTTGGCAGGCGCCACTGGATCCCGAGTGGGAGT 1906
Db 3716 AGACATGGGAGACATGGTGACCGAGTACTTGGCAAGCCACTGGATCCCTGAGTGGGAGT 3775
QY 1907 TCGTGAACACCCCCCTCGTGAAGCTGTGTACCAAGCTGGAGAGGAGGCCATCATCG 1966
Db 3776 TCGTGAACACCCCTCGTGTGAAGCTGTGTATCAGCTGGAGAGGAAACCCATCGTG 3835
QY 1967 GCGCCGAGACCTTCTAGCTGGAGCGCGCCCAACCGCAGACCAAGATCGGCAAGCGG 2026
Db 3836 GAGCAGAGACCTTCTAGCTGGATGGGCGACGCCAACAGAGGAGACCAAGCTGGCAAGSCAG 3895
QY 2027 GCTACGTGACCGACCGGGCGGCGAGAGATCGTGAGCCCTGACCGAGACCAACACAGA 2086
Db 3896 GCTACGTGACCAACCGGAGGACGACAGAAAGTGTGTGACCTGACTGACACACCAACACAGA 3955
QY 2087 AGACCGAGCTGACGGCCCTGAGCTGGCCCTGCAGGACGCGCAGGAGTGAACATCG 2146
Db 3956 AGACTGAGCTGCAAGCCACTACTAGCTCTGCAAGACACCGACTGGAAGTGAACATCG 4015
QY 2147 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCCGACAGAGGAGCGG 2206
Db 4016 TGACAGACTCACAGTACGCACTGGGCATCATCCAAGCACACACAGACCAATCCGAGTCAG 4075
QY 2207 AGCTGGTGAACAGATCATCGACGCTGATCAAGAGGAGAGGTGTACTCTGAGCTGGG 2266
Db 4076 AGCTGGTGAACAGATCATCGACGCTGATCAAGAGGAGAGGTGTACTCTGGCATGGG 4135
QY 2267 TGCCCGCCCAAGAGGCGATCGCGGCAACGACGAGATCGACAAAGCTGGTGACGAAGGCA 2326
Db 4136 TACCAGCACAAAGGAAATGGAGGAAATGAACAAGTAGATATAATAGTCAAGTCTGGGA 4195
QY 2327 TCCGCAAGTGTCTTCTCTGAGCGGCATCGAT 2358
Db 4196 TCCGGAAGGTGCTTCTCTGAGCGGCATCGAT 4227

RESULT 3

US-09-872-733-1
; Sequence 1, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, HIV GAG AND
; TITLE OF INVENTION: HIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, HIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4338
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Gag/Pol gene
US-09-872-733-1

Query Match 61.2%; Score 1503.8; DB 10; Length 4338;
Best Local Similarity 79.8%; Pred. No. 2.6e-239;
Matches 1896; Conservative 0; Mismatches 447; Indels 34; Gaps 9;

QY 14 TGGCCGAGGCGCATGAGCCAGGCCACACAGC--GCCACATCTCTGATGACGGCAGCAACT 70
Db 1085 TGGCCGAGGCGCATGAGCCAGGTGACGAACCTCGGCGACCATAAATGATGACAGAGGCAACT 1144
QY 71 TCRAAGGCCCCCAAGCGCATCATCAAGTGTCTCAACTCGGCAAGGAGGCCACATCGCCC 130
Db 1145 TCCGGAACACGAGGAAGATCGTCAAGTGTCTCAATTTGGCAAGAGGCGCACACCGCCA 1204
QY 131 GCAACTCGCCGCGCCCGCCCAAGAAAGGCTGTGGAAGTGGCGCAAGGAGGCCACACAGA 190
Db 1205 GGAAGTGGCCGCGCCCGCCGGAAGAGGCGTGTGGAATGTGGAAGAGGAGCACCAAA 1264
QY 191 TGAAGGACTGCACGAGCGCCAGGCCAACTTCTTCGCGAGGAGCTGGCCTTCCCCCAGG 250
Db 1265 TGAAGAGATTGACTGAGAGACAGGCTAA-TTTTATTAGGAAGATCTGGCCTTCTCAACG 1323
QY 251 GCAAGGCCCCGAGTTCGCCAGCGAGCAGAACCGCCGAGGCGCGCGAGCGCCACCGCCGAGC 310
Db 1324 GGAAGGCCCCAGGAAATTTCTTCAGAGCAGACAGGCCAACAGCCGCCACCCAGAAAGAGC 1383
QY 311 TGCAAGTGTCCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGAGCGCCAGCGGCA 364
Db 1384 TTCAGGTCTGGGTGAGAGCAACAACCTCCGCCCTCAGAAGCAGGAGCGCATAGACAAGAA 1443
QY 365 -----CCCTGAACCTTCCCGCAGATCAACCTGTGGCAGCGCCCTGTGTGAGCATCAAG 418
Db 1444 CTGTATCTTTAACTTCTCCCTCAGATCACTCTTTGGCAACGACCCCTGTGTGATAGGA 1503
QY 419 TGGCGCGCCAGATCAAGGAGGCGCTGTGTGACACCGCGCGCGAGCAGACCGTGTGAGG 478
Db 1504 TCGGGGGCAACTCAAGAAAGCGCTGCTCGATACAGGAGCAGATGATACAGTATTAGAAG 1563
QY 479 AGATGAGCCCTTCCCGGCAAGTGGAAAGCCCAAGATGATCGGCGGCATCGGGGCTTTCATCA 538
Db 1564 AAATGAGTTTTCGCGAGGAAGATGAAACCAAAATGATAGGGGGATCGGGGCTTTCATCA 1623
QY 539 AGGTGCCCGCAGTACGACGACATCTGTATCGAGATCTCGGCGAAGAGGCCATCGGACCG 598
Db 1624 AGGTGAGCGAGTACGACGACATCTATAGAAATCTGTGGACATAAAGCTATAGGTACAG 1683
QY 599 TGTGTATCGCCCGCCACC-----CCGCTGAACATCATCGCCGCAACATGCTGACCCAGCTG 654
Db 1684 TATTAGTAGGACCTTACCTACACCTGTCAACATTAATGGAAAGAAATCTGTTGACCCAGATC 1743
QY 655 GGCTGACCCCTGAACCTTCCCGCATCAGCCCATCGAGACCGTGCCTGTAAGCTGAAGCCC 714
Db 1744 GGCTGACCTTGAACCTTCCCGCATCAGCCCTATTGAGACGGTGCCTGTAAGCTGAAGCCG 1803
QY 715 GGAATGACGGCCCGCAAGGTGAAGCAGTGGCCCTGTACCGAGGAGAGATCAAGGCCCTG 774
Db 1804 GGAATGACGGCCCGCAAGGTGAAGCAGTGGCCCTATTGACGAAAGAGAGATCAAGGCCCTA 1863
QY 775 ACCGCCATCTCGGAGGAGATGGAGAGGAGGCGCAAGATCACCAAGATCGGCCCGAGAAC 834
Db 1864 GTCGAATCTGTACAGAGATGGAGAGGAGGAGGAGATCAGCAAGATCGGGCCTTGAGAAC 1923
QY 835 CCCTACAACACCCCGCTGTTCGGCATCAAGAAGAGGAGCAGCACCAAGTGGCGCAAGCTG 894
Db 1924 CCCTACAACACTCCAGTCTTCGGCAATCAAGAAGAGCAGTACCAAGTGGAGAAAGCTG 1983
QY 895 GTGAGCTTCCGAGCTGAACAAGCGCACCCAGACTTCTGGGAGGTGCAAGTGGGCGATC 954
Db 1984 GTGAGCTTCAAGAGCTGAACAAGAGAACTCAGGACTTCTGGGAAAGTTCAGCTGGGCATC 2043
QY 955 CCCCACCCCGCCCGCTGAAGAAGAGAGAGCGGTGACCTGTGGACGTGGGCGAGGCC 1014
Db 2044 CCATCCCGCTGGGTGAAGAAGAGAGAGTCACTGACAGTGTGGATGTGGGTGATGCC 2103

QY 1015 TACTTACGCTGGCCCTGGAGGAGCTTCGCAAGTACACCGCTTCCACTCCCGCAG 1074
DB 1016 TACTTACGCTGGCCCTGGAGGAGCTTCGCAAGTACACCGCTTCCACTCCCGCAG 1075
DB 2104 TACTTCCGCTTCCTTGGAGGAGCTTCAGGAAGTACACTGCTCCCTCAGGATACCTAGC 2163
QY 1075 ATCAACACGAGACCCCGGCGATCCGCTACCGATGAACTGCTGCTCCCGCAGGCTGGAG 1134
DB 2164 ATCAACACGAGACCCCGGCGATCCGCTACCGATGAACTGCTGCTCCCGCAGGCTGGAG 2223
QY 1135 GGCAGCCCGGAGCTTCTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCGCGCCCGC 1194
DB 2224 GGATCACCACCCATCTTCAAGACAGCATGACCAAGATCCTGGAGCCCTTCGCGAGCAA 2283
QY 1195 AACCCCGAGATCGTATCTTACCA-----GGCCCCCTGTACCTGGGCGAGCAGCTGGAG 1248
DB 2284 AACCCAGACATCGTATCTTATCAGTACATGGACGACCTCTACGTAGGAAAGTACCTGGAG 2343
QY 1249 ATC-GGCCACGACCGGCCCAAGATCGAGGAGCTGGCGAGCAGCTGCTGGCTGGGGCTT 1307
DB 2344 ATCGGGGCGACGAGGACCAAGATCGAGGAGCTTGAGACAGCATCTGTTGAGGTGGGGACT 2403
QY 1308 CACCACCCCGCAAGAAGCACCAGAGGAGCCCTTCTTCCGCCAT-----CGAGCT 1361
DB 2404 GACCACACAGACAGAGGACCAAGAGGAACTCTCTCTCTGTGATGGCTACCAACT 2463
QY 1362 GCACCCCGCAAGTGGACCGCTGCAGCCCATCGAGCTGCCCGAGAGGAGCTGGACCGT 1421
DB 2464 GCATCTGCAAGTGGACAGTGCAGCCCATCGTGTGCTGAGAGGAGCAGCTGGACTGT 2523
QY 1422 GAAGGACATCCAGAGCTGTGGGCAAGCTGAAGTGAAGTGGGCGCAGCAGATCTACCCGGCAT 1481
DB 2524 GAAGGACATCCAGAGCTGTGGGCAAGTGAAGTGAAGTGGGCGCAGCAGATCTACCCGGCAT 2583
QY 1482 CAAGGTCGCGCAGCTGTGCAAGTGTGCGCGCGCCAGGCGCTCAGCGACATCTGCC 1541
DB 2584 CAAGTGTAGCAGCTGTGCAAGTGTGCGCGCGCCAGGCGCTCAGCGACATCTGCC 2643
QY 1542 CCTGACCGAGGCGCGAGCTGGAGCTGGCGCGAGACCGGAGATCCTGCGGAGCGCGT 1601
DB 2644 ACTGACAGAGGAGCAGAGTAGAATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2703
QY 1602 GCAGGCGGTACTACGACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1661
DB 2704 ACATGGAGTGTACTACGACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2763
QY 1662 CGACAGTGGACCTTACAGATCTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1721
DB 2764 AGCCATGAGCTTACCAATCTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2823
QY 1722 CGCCAGATCGGACCGCCCGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1781
DB 2824 CGCAAGGATCAGGGGTGCGCCACACCAACGATGTGAAGCAGCTGACAGAGGAGGAGGAG 2883
QY 1782 GATCGCATGGAGAGATCTGTATCTGGGCGAGAGCCCGAAGTTCGCGCTGCCCATCCA 1841
DB 2884 GATCACACAGAGAGATCTGTATCTGGGCGAGAGCTCCCAAGTTCGCGCTGCCCATCCA 2943
QY 1842 GAAGGAGACCTGGAGACCTTGTGGACCGACTACTTGGAGGAGGAGGAGGAGGAGGAGGAG 1901
DB 2944 GAAGGAGACATGGAGAGATGTGTGGACCGAGTACTGGCAGGAGGAGGAGGAGGAGGAG 3003
QY 1902 GGAGTGTGTGAACACCCCGCCCTTGTGTGAAGCTGTGTGTACAGCTGTGGAGGAGGAGGAG 1961
DB 3004 GGAGTGTGTGAACACCCCGCCCTTGTGTGAAGCTGTGTGTATCAGCTGTGGAGGAGGAG 3063
QY 1962 CATCGGCGCGAGACCTTCTACGTGAGCGGCGCGCCCAACCGGAGAGACCAAGATCGGCAA 2021
DB 3064 CGTGGGAGCAGAGACCTTCTACGTGAGCGGCGCGCCCAACCGGAGAGACCAAGATCGGCAA 3123
QY 2022 GGCCTGTGTGACCGAGCGGCG 2081
DB 3124 GCGAGGCTAGTGACCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3183
QY 2082 CCAGAAGACCGAGCTGACGGGCAATCCAGCTGGCCCTTGAGGAGGAGGAGGAGGAGGAG 2141

DB 3184 CCAGAAGACTGAGCTGCRAGCCATCTACTAGCTCTGCAAGACGCGGACTGGAGTGAA 3243
QY 2142 CATCTGACGACGACGAGCTAGTCCCTGGGATCATCTCCAGGCGCCAGCCGACAGAGGGA 2201
DB 3244 CATCTGACGAGCTACAGTACG-CATGGCATCATCCCAAGCACAACAGACCAATCCGA 3302
QY 2202 GAGCGAGCTGGTGAACAGATCATCGAGCAGTGTATCAAGAGGAGGAGGAGGAGTGTACTGAG 2261
DB 3303 GTACAGCTGGTGAACAGATCATCGAGCAGTGTATCAAGAGGAGGAGGAGGAGTGTACTGCGC 3362
QY 2262 CTGGTGGCCGCCCAAGAGGAGCTCGGCGCAACGAGCAGATCGACAAAGCTGGTGAGCAA 2321
DB 3363 ATGGGTACCAGCAGCACAAGAGAAATGGAGGAAATGAACAAGTAGATAAATAGTCAAGTGC 3422
QY 2322 GGGCATCGCAAGGTGCTTCTCTGGACGGCATCGAT 2358
DB 3423 TGGGATCCGGAAGGTGCTGTTCTCTGGACGGATCGAT 3459

RESULT 4
US-09-872-733-3
; Sequence 3, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287051 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872.733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Pol gene
US-09-872-733-3

Query Match 46.3%; Score 1137; DB 10; Length 2467;
Best Local Similarity 83.4%; Pred. No. 5e-179;
Matches 1319; Conservative 0; Mismatches 250; Indels 12; Gaps 2;

QY 790 GAGATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 849
DB 7 GAGATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 66
QY 850 GTGTTCGCGCATCAACAAGAAGGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 909
DB 67 GTCTTCGCAATCAAGAAGAGGAGCAGTACCAAGTGGAGGAGGAGGAGGAGGAGGAGGAG 126
QY 910 CTGAACAAGCGACCCAGGAGCTTCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 969
DB 127 CTGAACAAGAGAACTCAGGAGCTTCTGGGAAGTTCAGCTGGGAGGAGGAGGAGGAGGAGGAG 186
QY 970 CTGAAGAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1029
DB 187 TTGAAGAAGAAGAAGTCAAGTGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGT 246
QY 1030 CTGGACGAGGAGCTTCCGCAAGTACACCCCTTACCATCCCGCAGCATCAACAAGAGAGACC 1089
DB 247 TTGGACGAGGAGCTTCAAGAGTACACTGCCCTTCAAGTACCTTCAAGTACCTTCAAGTAC 306
QY 1090 CCGGCGATCCGCTACCAAGTACCAAGTGTGCTGCCCGAGGAGGAGGAGGAGGAGGAGGAGGAG 1149
DB 307 CCAGGCGATCCGCTACCAAGTGTGCTGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 366

; CURRENT FILING DATE: 2000-12-14
 ; PRIOR APPLICATION NUMBER: 09/117,217
 ; PRIOR FILING DATE: 1998-07-24
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 2601
 ; TYPE: DNA
 ; ORGANISM: HIV-HXB2
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (334)..(489)
 ; OTHER INFORMATION: gag P6 (52 AA)
 US-09-735-487-9

Query Match 44.0%; Score 1081.6; DB 10; Length 2601;
 Best Local Similarity 67.9%; Pred. No. 6.4e-170;
 Matches 1610; Conservative 0; Mismatches 734; Indels 28; Gaps 6;

QY	14	TGCGCGAGGCGCATGAGCGAGCGCACCA--GCGCCAACATCTCTGATGCGAGCGAGCAACT	70
DB	77	TGCTGAAGCAATGAGCGAAGTAACAATTCAGTACCATAATGTCAGAGGCAATT	136
QY	71	TCAAGGCGCCCAAGCGCATCATCAAGTCTTCACTGCGCAAGGAGGCGGCACATCGCCC	130
DB	137	TTAGGAACCAAGAAGATTTGTTAAGTGTTCATTTGTGCGCAAGAGGCGGCACACAGCCA	196
QY	131	GCAACTGCGCGCGCCCGCCCAAGAGGCTGCTGGAAGTCGCGCAAGGAGGCGCACAGA	190
DB	197	GAAATTCAGGGGCGCTTAGGAAGAGGCTGTGGAATGTGGAAGGAGGAGGACCAAA	256
QY	191	TGAAGGACTCACCGAGCGCGAGCGCAACTTCTTCCGCGAGGACCTGGCCCTTCCGCCAGG	250
DB	257	TGAAGATTCTAGAGAGACAGCTAA-TTTTATGGAAGATCTGGCCCTTCTTACAAG	315
QY	251	GCAAGGCGCGGATGTCCTCCAGGAGGAGAGCAACCGCGCCACAGCCCGCAGCGGAGC	310
DB	316	GGAAGGCGAGGGAATTTCTTCAAGAGAGAGACAGAGCCACAGCCCGCAGAGAGAGC	375
QY	311	TGCAGGTGCGCGG-----CGACAACCGCGAGCGAGCGCGCGCGAGCGCGCAGGGCA	364
DB	376	TTGAGTCTGGGTAGAGACACAACACTCCCGCCAGAGGAGGCGCGATAGACAGGAA	435
QY	365	-----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAG	418
DB	436	CTGTATCTTTAACTTCCCTCAGTCACTCTTTGGCAACGACCCCTCGTCACAATAAGA	495
QY	419	TGGCGCGCAGATCAAGGAGCGCTGCTGGACACCGCGCGCGAGGACCGTGTGGAGG	478
DB	496	TAGGGGGCAACTAAGGAAAGCTTATTAGATACAGGAGCAGATGATACAGTATTAGAG	555
QY	479	AGATGAGCCTGCGCGGCAAGTGAAGCGCAAGATGATGCGCGCATCGCGCGGTTCATCA	538
DB	556	AATGAGTTTGCCAGGAGATGGAACCAAAATATAGTAGGGGAATTGGAGTTTATCA	615
QY	539	AGGTGCGCCAGTACGACAGATCCTGATCGAGATCTGCGCAAGAGGCGCATCGGACCG	598
DB	616	AAGTAAGACAGTATGATCAGATCTATAGAAATCTGTGACATATAAGCTATAGGTACAG	675
QY	599	TGCTGATCGCGCCCGCCCGTGAACATCATCGGCCCGCAACATGCTGACCCAGCTGGGCT	658
DB	676	TATTAGTAGGACCTTACACCTGTCAACATAATTTGGAAGAAATCTGTTGACTCAGATTGGTT	735
QY	659	GCACCTGAACTTCCCATCAGCCCATCGAGACCGGTGCGCGTGAAGCTGAAGCGCGCA	718
DB	736	GCACCTTAAATTTTCCATTTAGCCCTATTGAGACTGTACCAGTAAATTTAAGCCAGGAA	795
QY	719	TGAGCGCCCAAGGTGAAGCAGTGGCCCTGACCGGAGGAGAGATCAAGGCCCTGACCG	778
DB	796	TGGATGCGCCCAAGGTATAACAATGGCCATTGACAGAGAGAAAAATAAAGCATTAGTAG	855
QY	779	CCATCTCGAGGAGATGGAGAGGAGGAGATCACCAGATCGGCCCGCGAGAACCCCT	838

DB	856	AAATTTTACAGAGATGGAAGAGAGCGAAAAATTTCAAAAAATTTGGCCCTGAAAAATCCAT	915
QY	839	ACAACACCCCGTGTTCGCCCATCAAGAGAGAGACACACCAAGTGTGCGCAAGCTGTGG	898
DB	916	ACAATCTCCAGTATTGGCCATTAAGAAAAAGACAGACTACTAATATGGAGAAAAATTAGTAG	975
QY	899	ACTTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCAATCCCC	958
DB	976	ATTTACAGAACTTAATAAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCAC	1035
QY	959	ACCCCGCGCGCTGAAGAAGAGCGTGCACCGTGTGGAGCTGGAGCTGGGCGACGCCCTACT	1018
DB	1036	ATCCCGCAGGCTTAAAAAGAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATT	1095
QY	1019	TCAGCGTCCCGCTGGAGGAGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA	1078
DB	1096	TTTCAGTTCCTTAGATGAAGACTTCAGGAAGTATATCTGCATTTACCATACCTAGTATAA	1155
QY	1079	ACAACAGAGACCCCGCATCCGCTACAGTACAACTGCTGCCCGAGGGGTGGAGGGCA	1138
DB	1156	ACAATGAGACACCGGATTAGATATCAGTACAATGTGCTTCCACAGGATGGAAGGAT	1215
QY	1139	GCCCCAGCATTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCCCGCAACC	1198
DB	1216	CACGACCAATTTCCAAAGTAGCATGACAAAAATCTTTAGAGCCCTTTTAGAAAAAATC	1275
QY	1199	CCGAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGGCGAGCGACCTGGAGATCG	1252
DB	1276	CAGACATGTTTCTATCATATCATATGATGATGTTGTATGATGATCTAGGAATAG	1335
QY	1253	GCAGACACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGTGGGGCTTCACCA	1312
DB	1336	GGCAGCATAGAAAAAATAGAGGAGCTGAGACAACTCTGTTGAGGTGGGGCTTACCA	1395
QY	1313	CCCCGACAGAGACCAAGAGGAGCCCGCTTCTTCCCTGCCCCAT-----CGAGCTGCACC	1366
DB	1396	CACGACAAAAAACATCAGAAAAAGCTCCATTTCTTGGATGGGTATTGAATCCATC	1455
QY	1367	CCGACAACTGAGCGCTGCAGCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG	1426
DB	1456	CTGATAATGGACAGTACAGCTATAGTGTGCCAGAAAAAGACAGCTGGACTGTCAATG	1515
QY	1427	ACATCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCGGCATCAAG	1486
DB	1516	ACATACAGAACTAGTGGGAAATTTGAATTTGGCAAGTCAAGATTTACCCAGGATTAAG	1575
QY	1487	TGGCGCAGCTGCAAGCTGCTGCGCGCGCAGAGGCGCTGACCGACATCGTGGCCCTGA	1546
DB	1576	TAAGGCAATTTATGTAAACTCTTAGAGGAACCAAGACCTAACAGAAAGTAAATACCACTAA	1635
QY	1547	CCGAGGAGCGCGCTGGAGCTGGCGGAGAACCGCGAGATCTCTCGCGAGCGCGTGCACG	1606
DB	1636	CAGNAGAGCAGAGCTAGAACTGCGCAAAACAGAGAGATTTCTAAAGAACCACTACATG	1695
QY	1607	GCGTGTACTACGACCCCGAGGAGCTGTGGCGCAGATCCAGAGAGGCGGCCACGAC	1666
DB	1696	GAGTGTATTATGACCATCAAAAGACTTAATAGCAGAAATACAGAGGCGGGCAAGGCC	1755
QY	1667	AGTGGACCTACCATCTACAGAGCGCTTCAAGAACCTTCAAGACCGGCGCAAGTACGCCA	1726
DB	1756	AATGGACATCAAAATTTATCAGAGGCCATTTAAAAATCTGAAAAACAGAAAAATATGCAA	1815
QY	1727	AGATGCGCAGCGCCACCAACAGCTGAAGCAGCTGACCGAGCGCGTGCAGAAAGATCG	1786
DB	1816	GAATGAGGGTGGCCACACTAATGATGTAAACAATTAACAGAGCGAGTGCAAAAATAA	1875
QY	1787	CCATGAGAGCATGCTGATCTGGGCGAAGACCCCAAGTTCCCGCTGCCCATCCAGAGG	1846
DB	1876	CCAGAGAAAGCATAGTAATATGGGAAAGACTCTCTAAATTTAAACTGCCCATACAAAGG	1935
QY	1847	AGACCTGGAGAGCTGCTGGACCGACTACTGGCAGGCGACCTGTGATCCCGAGTGGAGT	1906
DB	1936	AAACATGGGAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCCTGAGTGGAGT	1995

QY	1907	TCGTGAACACCCCCCTCGTGTGAAGCTGTGGTACCGAGTGGAGAAGGCCCATCATCG	1966
Db	1996	TTGTTTAATACCCCTCCCTTAGTGAATTAATGTCACGATTAGAGAAAAGAACCCCATAGTAG	2055
QY	1967	GCGCCGAGACCTTCTACGTGGAGCGCGCCCAACCGCGAGACCAAGATCGGGAAGCCG	2026
Db	2056	GACCAGAAACCTTCTATGTAGATGGCGCAGCTACAGGGAGACTTAATTTAGRAAAGCAG	2115
QY	2027	GCTACGTGACCGACCGGGCGGCAAGAAGATCGTGAGCCTGACCGAGACCACCAACCCAGA	2086
Db	2116	GATATCTTACTAATAGAGGAAGACAAAAAGTTGTACCCCTAACTGCACACAACAAATCAGA	2175
QY	2087	AGACCCAGCTGCAGGCCATCAGCTGCGCCCTGCGAGGACAGCGGCAGGAGGTGAACATCG	2146
Db	2176	AGACTGAGTTTCAAGCAATTTATCTAGCTTTTGCAGGATTCGGGATTAGAAGTAAACATAG	2235
QY	2147	TGACCCAGACGCGAGTACGCCCTGGCGCATCATCCAGGCCACCCCGACAAGAGCGGAGCG	2206
Db	2236	TAACAGACTCACAATATGCATTAGGAATCATTCACGCACACACAGATCAAGTGAATCAG	2295
QY	2207	AGCTGGTGAACAGATCATCGAGCAGCTGTATCAGAAGAGGAAGGTGTACTCGAGCTGGG	2266
Db	2296	AGTTAGTCAATCAAAATAATAGACGAGTTTAATAAAAAGGAAGGTCTATCTGGCATGGG	2355
QY	2267	TGCCGCCCCACAAGGGCATCGCGGCAACGAGCAGATCGACAACCTCGTGAGCAAGGGCA	2326
Db	2356	TACCACGACACAAGGAATTGGAGGAATGACACAGTAGATAAATTTAGTCAGTCTGTGGAA	2415
QY	2327	TCGCGAAGTGTGTTTCCTGGACGCCATCGAT	2358
Db	2416	TCAGGAAAGTACTATTTTATAGATGAATAGAT	2447

RESULT 7

```

US-09-735-487-11
; Sequence 11, Application US/09735487
; Patent No. US20020042679A1
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/735,487
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/117,217
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (453)..(749)
; OTHER INFORMATION: Protease
US-09-735-487-11

```

	Query Match	44.0%	Score 1081.6;	DB 10;	Length 2601;
	Best Local Similarity	67.9%	Pred. No. 6.4e-170;		
	Matches 1610;	Conservative 0;	Mismatches 734;	Indels 28;	Gaps 6;
QY	14	TGCGCGAGCCATGAGCCAGCGCCACCA	---	GCGCCAACATCTTGATGCGAGCGAGCAACT	70
Db	77	TGCGTGAAGCAATGAGCCAAGTACAAATTCAGCTACCATAATGATGCAGAGGCGCAATT			136
QY	71	TCAGGGGCCCCAAGCGCATCATCAAGTGCTTCAACTCGGCCAAGAGGGGCCACATGCCCC			130
Db	137	TTAGGAACCCAAAGAAGAAGATTGTTAAGTGTTTTCAATTGTGCGAAGAAGGGGCACACGCCA			196

Db 556 AATAGAGTTTCCAGGAAGATGAAACCAAAATGATAGGGAAATTTGGAGTTTATCA 615
QY 539 AGGTGCCAGTAGACAGATCTGATCAGATCTCGCGCAGAAGAGCCATCGGCAACG 598
Db 616 AAGTAGACAGTAGATGATCAGATCTCATAGAAATCTGTGACATAGAGTATAGGTACAG 675
QY 599 TGTGTATCGGCGCCACCCCGTGAACATCATCGCGCGCAACATCTGTGACCGAGCTGGGT 658
Db 676 TATTAGTAGACCTACACCTGTCAACATATTGGAAGAAATCTGTGACTCAGATTGGTT 735
QY 659 GCACCCCTGAATCTCCCATCAGCCCATCAGACCCATGAGACCCGTGGAAGCTGGAAGCCCGGA 718
Db 736 GCATTTTAAATTTTCCCATTAGCCCTATTGAGACTGTACAGTAAATTAAGGCCAGGA 795
QY 719 TGGAGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 796 TGGATGCCCAAGATTAAACATAGGCCATTGACAGCAAAAAATAAAGCATTTAGTAG 855
QY 779 CCATCTCGAGGAGATGAGAAGAGGGCAAGATCACAGATCGGCCCGCCGAGAACCCCT 838
Db 856 AATTTCTACAGAGATGAAAGAGAGCGAAATTTCAAAATTTGGCCCTGAAATCCAT 915
QY 839 ACAACACCCCGTGTTCGCCATCAGAAGAGACAGACCAAGTGGCGCAAGCTGGTG 898
Db 916 ACAATACTCCAGTATTTCGCATTAAGAAAAAAGACAGTACTAAATGGAGAAATTTAGTAG 975
QY 899 ACTTCCGCGAGCTGAACAAGCGCACCGAGCTTCTGGAGGTGCAGCTGGGCATCCCC 958
Db 976 ATTTTCAGAGACTTAAAGAGACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCAC 1035
QY 959 ACCCGCGCGCTGAAGAAGAAAGAGCGTGACCGTGTGGAGCTGGCGAGCGCCTACT 1018
Db 1036 ATCCGCGAGCTTAAAGAAAGAAATCACTACAGTACTGGATGTGGTGTGCATAT 1095
QY 1019 TCAGCTGCCCCGTGGAGGAGACTTCGGCAAGTACACCGCTTACACATCCCCAGCATCA 1078
Db 1096 TTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATCTGCTATTACCATACCTAGTATAA 1155
QY 1079 ACAAGAGACCCCGGATCCGTACAGTACAACTGCTGCCCGAGGCTGGAAGGCA 1138
Db 1156 ACAATGAGACACAGGATTAGATATCAGTACAAATGTCTTCACAGGATGAAAGGAT 1215
QY 1139 GCCCGAGCATCTCCAGAGCAGATGACCAAGATCTGGAGCCCTTCGCGCGCCGCAAC 1198
Db 1216 CACGAGCAATATCCAAAGTAGCATGACAAAAATCTTAGAGCCTTTTAGAARACAAATC 1275
QY 1199 CCGAGATCGTGTACCA-----GGCCCCCTGTACGTGGCGAGCGACTGGAGTCG 1252
Db 1276 CAGACATAGTTATCTATCAATACATGATGATTTGTATGTATGAGTCTGACTTAGAATAAG 1335
QY 1253 GCCAGACCGCGCAAGATCGAGAGTGGCAAGCAGCTGCGCTGGGCTTCACCA 1312
Db 1336 GGCAGCATAGACAAAAATAGAGAGCTGAGACACATCTGTTGAGGTGGGACTTTACCA 1395
QY 1313 CCCCCGAGAAAGCACCAGAGAGCGCCCTTCTGCCCCAT-----CGAGCTGCACC 1366
Db 1396 CACGAGACAAAAACATCAGAAAGAACCTCCATTCTCTTGGATGGGTATGAACTCCATC 1455
QY 1367 CCGAAGTGGAGCGTGCAGCCCATGAGTGGCGGAGAGAGAGTGGACCGTGAACG 1426
Db 1456 CTGATAATGGAGTAGTACAGCCTATAGTGTGCGCAAAAAGACAGCTGACTGTCAATG 1515
QY 1427 ACATCCAGAAGCTGTGGGCAAGCTGAACCTGGCCAGCCAGATCTACCCCGCATCAAG 1486
Db 1516 ACATACAGAAGTGTAGTGGGAAATTAATTTGGCAAGTCAAGTTTACCCAGGATTAAG 1575
QY 1487 TGGGCCAGCTGTGAAGCTGCTGGGGCGCCCAAGCCCTGACCGACATCGTCCGCCCTGA 1546
Db 1576 TAAGGCAATTATGTAACTCTTAGAGGAACCAAGCACTAACAGAGTAATACCACTAA 1635
QY 1547 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTCGCGCGCCGCTGACG 1606

Db 1636 CAGAAGAAGCAGAGCTAGAACTGGCAGAAAAACAGAGAGATTCTAAAGAACCACTACATG 1695
QY 1607 GCGTGTACTACGACCCAGCAGGAGACCTGGTGGCGAGATCCAGAGCAGGCGCACGACC 1666
Db 1696 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAAGAAATACAGAAAGCGGGCANGGCC 1755
QY 1667 AGTGGACCTTACCAGATCTACCAGGAGCCCTTCAAGAACTTGAAGACCGGCAAGTACGCCA 1726
Db 1756 AATGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAACACAGAAAATATGCAA 1815
QY 1727 AGATCGGACCCCGCCACACCAACGACGTGAAGCAGCTGACCGAGCCCTGAGAGATCG 1786
Db 1816 GAATGAGGGGTGCCACACTATATGATTAACAACTTAAACAGAGCAGTGCACAAAAATAA 1875
QY 1787 CCATGGAGAGCTGATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAAAG 1846
Db 1876 CCACAGAAAGCATAGTAATATGGGAAAGACTTCTTAATTTAACTGCCCATACAAAAG 1935
QY 1847 AGACCTGGAGACCTGGTGGACCGGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1906
Db 1936 AAACATGGAAACATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTTGAAGTGGAGT 1995
QY 1907 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACCACTGGAGAGAGAGCCCATCATCG 1966
Db 1996 TTGTTAATACCCCTCCCTTAGTGAATTTATGTTACCACTTAGAGAAAGAACCCCATAGTAG 2055
QY 1967 GCGCGAGAGCTTCTACGTGGAGCGCGCCCAACCGGAGACCAAGATCGCAAGGCG 2026
Db 2056 GAGCAGAAACCTTCTATGTAGTGGGAGCTTAACAGGAGACTTAATTTAGAAAGCAG 2115
QY 2027 GCTAGCTGACCGACCGGGCGGCAAGAGATCGTGAGCTGACCCGAGACCCACCAACAGA 2086
Db 2116 GATATGTTACTAATAGAGAAAGAAAGTTGTCAACCTTACTGACACACAAATCAGA 2175
QY 2087 AGACGAGCTGCGAGCGCATCTCAGCTGGCCCTGCGAGGACGCGGAGAGGTGAACATCG 2146
Db 2176 AGACTGAGTTTACAAGCAATTTATCTAGCTTTGAGGATTCGGGATTAAGAGTTAAACATAG 2235
QY 2147 TCAGCGACCGCAGTACGCCCTGGGCATCTCCAGCGCCAGCCCGCCAGAGAGCGG 2206
Db 2236 TAACAGACTCACAATATGCAATTTAGGAATCATTTCAAGCAACACAGCATCAAGTGAATCAG 2295
QY 2207 AGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGG 2266
Db 2296 AGTTAGTCAATCAATAATAGAGCAGTTAATAAAGAAAGAAAGTCTATCTGGCATGG 2355
QY 2267 TCCCGCGCCACAAGGCGATCGCGGCGCAGAGCAGATCGACAAGCTGGTGAAGAGGCA 2326
Db 2356 TACCAGCACACAAAGGAATTTGAGGAAATGACAAAGTAGATAAATAGTCTAGTCTGGAA 2415
QY 2327 TCCGCAAGGTGCTGTTCTCTGGACGCGCATCGAT 2358
Db 2416 TCAGGAAGTACTATTTTAGATGAATAGAT 2447

RESULT 9
US-09-999-183-1
; Sequence 1, Application US/09999183
; Patent No. US20020147169A1
; GENERAL INFORMATION:
; APPLICANT: MITROPHANOUS, et al
; TITLE OF INVENTION: In Vivo Selection Method
; FILE REFERENCE: 674523-2009
; CURRENT APPLICATION NUMBER: US/09/999,183
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: PCT/GB00/02136
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 9912965.2
; PRIOR FILING DATE: 1999-06-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: SeqWin99
; SEQ ID NO 1
; LENGTH: 4307

; TYPE: DNA									
; ORGANISM: Human immunodeficiency virus type I									
; -09-999-183-1									
Query Match 44.0%; Score 1081.6; DB 10; Length 4307;									
Best Local Similarity 67.9%; Pred. No. 6.6e-170;									
Matches 1610; Conservative 0; Mismatches 734; Indels 28; Gaps 6;									
QY	14	TGGCCGAGGCATGAGCAGGCACCA--GCGCCAACTCTCTGATCAGCGGCAACT	70						
Db	1088	TGGCTGAAGCAATGAGCCAGTACAAATTCAGTACCTACCAATGATCGAGAGGCAATT	1147						
QY	71	TCAAGGCCCCCAAGCGCATCATCAAGTGTCTCAACTCGGCAAGGAGGCCACATCGCCC	130						
Db	1148	TTAGGAACCAAGAAAGATTTTAAAGTGTTCATTTGCGAAAGAGGCGACACAGCCA	1207						
QY	131	GCAACTCGCGCGCCCGCCGCAAGNAGGCTGCTGGAAGTGGCGAAGGAGGCCACACGA	190						
Db	1208	GAATTCGAGGGCCCTTAGGAAAGAGGCTGTGGAAATGTGAAAGGAAGGACACCAAA	1267						
QY	191	TGAAGGACTGCACGCGCGCAGGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCAGG	250						
Db	1268	TGAAGATTGCTAGAGACAGGCTAA-TTTTATAGGGAAGATCTGGCCTTCTACAG	1326						
QY	251	GCRAAGGCCCGGAGTTCCCGCAGCGAGCAGAAACCGCGCCAAAGCCCAACGCGGAGC	310						
Db	1327	GGRAAGGCCAGGGAATTTCTTTCAGAGCAGACAGACAGCCACAGCCGCCACAGAGAGC	1386						
QY	311	TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGCGCGCCAGGGCA	364						
Db	1387	TTCAGGTCTGGGTAGAGACAACAACCTTCCCCCTCAGAAGCAGGAGCGGATAGACAAGAA	1446						
QY	365	-----CCCTGAACTTCCCCAGATCACCCCTGTGGCAGCGCCCTTGTGAGCATCAAG	418						
Db	1447	CTGTATCTTTAACTTCCCTCAGTCACTCTTGGCAACGACCCCTCGTCACATAAGA	1506						
QY	419	TGGCGGCCAGATCAAGGAGGCGCTGTGTGACACCGCGCGCGCAGCACCCTGTGAGG	478						
Db	1507	TAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGA	1566						
QY	479	AGATGAGCCTGCCCGCAAGTGGAGCCCAAGATGATCGCGGCATCGCGCGCTTCATCA	538						
Db	1567	AAATGAGTTTCCACAGNAGATGAAACCAAAATGATAGGGGAATTTGAGGTTTATCA	1626						
QY	539	AGTGCGCCAGTAGCAGCAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG	598						
Db	1627	AAGTAAGACAGTATGATCAGATCTCATAGAAATCTGTGCACATAAAGCTATAGGTAC	1686						
QY	599	TGCTGATCGCCCCACCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT	658						
Db	1687	TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTT	1746						
QY	659	GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCGGCA	718						
Db	1747	GCACTTAAATTTTCCCATTTAGCCCTATTGAGACTGTACAGTAAATTAAGCCAGGAA	1806						
QY	719	TGACGSCCCCAAGGTGAAGCAGTGGCCCTGACCAGGAGAGAAGATCAAGGCCCTGACCG	778						
Db	1807	TGGATGSCCCCAAAAGTTAAACATGGCCATTGACAGAGAAATAAAGACATTAGTAG	1866						
QY	779	CCATCTCGGAGGATGGAGAGGAGGCAAGATCAACAAGATCGGCGCCCGAGAACCCCT	838						
Db	1867	AAATTTGTACAGAGATGGAAAGAGGAGGAAATTTCAAAATTTGGGCCCTGAAATCCAT	1926						
QY	839	ACAACACCCCGGTTCGCCCATCAAGAAGAGCAGCAGCAAGTGGCGCAAGCTGGTGG	898						
Db	1927	ACAATCTCAGTATTGGCCATAGAAGAAAGACAGTACTAAATGGAGAGAAATTAGTAG	1986						
QY	899	ACTTCCGCGAGCTGAACAAGCAGCCAGGACTTCTGGGAGGTGCAAGTGGGCTCCCC	958						
Db	1987	ATTTGACAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACC	2046						
QY	959	ACCCCGCGCCTGAAGAGAAGAGCGGTGACCGTGTGGACGTGGCGGACGCCTACT	1018						
Db	2047	ATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGTGCATAT	2106						
QY	1019	TCAGCGTCCCTTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA	1078						
Db	2107	TTTTCAGTCTCCCTTAGATGAAGACTTCAGGAAGTATATGCAATTTACCATCTAGTATA	2166						
QY	1079	ACAACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGGAAGGCA	1138						
Db	2167	ACAATGAGACACACAGGATTAGATATCAGTACATGTGCTTCCACAGGGATGGAAGGAT	2226						
QY	1139	GCCTCAGCATTTCCAGAGCAGCATGACCAAGATCTTGGAGCCCTTCGCGCCCGCAACC	1198						
Db	2227	CACCAGCAATATTCCAAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAACAAATC	2286						
QY	1199	CCGAGATCGTATCTACCA-----GCCCCCTGTGTAGTGGGCGAGGACCTGAGATCG	1252						
Db	2287	CAGACATAGTATTCTATCAATACATGGATGATTGTGTAGTAGTCTGACTTAGAAATAG	2346						
QY	1253	GCCAGCACCCCGCCCAAGATCGAGGAGCTGGCAAGCACCTGCTGCGTGGGGCTTCACCA	1312						
Db	2347	GGCAGCATAGACAAAAATAGAGGAGCTGAGACAACATCTGTGAGTGGGACTTACCA	2406						
QY	1313	CCCCCGACAAGAACACCAAGAGGAGCCCTTCTTGCCTAT-----CGAGCTGCACC	1366						
Db	2407	CACCAGCAAAAAACATCAGAAAGAACCTCCATTCTTGGATGGTATTGAATCCCATC	2466						
QY	1367	CCGACAAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAGGAGAGCTGGACCTGAACG	1426						
Db	2467	CTGATAAATGGACAGTACAGCCTATAGTGTGCCAGAAAAAGACAGCTGGACTCTCAATG	2526						
QY	1427	ACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCGCCAGCATCTACCCGGCATCAAG	1486						
Db	2527	ACATACAGAGTGTAGTGGGAAATTTGAATTTGGGCAAGTCAGATTTACCCAGGATTAAAG	2586						
QY	1487	TGGCCAGCTGTGCAAGCTGCTGCGCGGCCCAAGGCCCTTGACCGACATCGTGCCCTGA	1546						
Db	2587	TAAGGCAATTTATGTAAACTCTTTAGAGGAACCAAGCACATAACAGAAAGTAAATCACCTAA	2646						
QY	1547	CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGAGATCCTCGCGAGGCCCTGTGCAGC	1606						
Db	2647	CAGAAGAGCAGAGCTAGAACTGGCAGAAAAAGAGAGATTTCTAAAAGACCATAGATG	2706						
QY	1607	CGTGTACTACGACCCCGCAGCAAGCCTGGTGGCGCGAGATCCAGAACGAGGCCACGACC	1666						
Db	2707	GAGTGTATTATGCCCATCAAAAGACTTAATAGCAGAAATACAGACGAGGGCAAGGCC	2766						
QY	1667	AGTGGACCTACAGATCTACAGAGGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA	1726						
Db	2767	AATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAAACAGAAAAATATGCA	2826						
QY	1727	AGATGCGCACCCCGCCCATCGACGCTGAAGCAGCTGACCGAGGCGCTGCAGAAATCG	1786						
Db	2827	GAATGAGGGTGGCCACACTAATGATGTAACAATTTAAGAGAGGCAATGCAAAAAATA	2886						
QY	1787	CCATGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCCGCTCCCATCCACAGAG	1846						
Db	2887	CCACAGAAACCATAGTAATATGGGAAAGACTCCTAAATTTAACTCCCATACAAAGG	2946						
QY	1847	AGACTTGGGAGACCTGGTGGACCGACTACTTGGCAGGCCACTTGGATCCCGAGTGGAGT	1906						
Db	2947	AAACATGGGAAGACATGGTGGACAGAGTATTGGCAAGCCACTGGATTCTCTGAGTGGGAGT	3006						
QY	1907	TCTGTGAACACCCCGCCCTGCTGNAAGCTGTGGTACCAGCTGGAGAGAGGCCCATCATCG	1966						
Db	3007	TTGTTAATACCCCTCCCTTTAGTGAATTTATGGTACCAGTTAGAGAAAGAACCCCATAGTAG	3066						
QY	1967	GCCTCCGAGACTTCTACGTGAGCGGCGCCGCAACCCGAGAGCAAGATCGGAAGGCCG	2026						
Db	3067	GAGCAGAACTTCTATGTAGTAGTGGGCAGCTAACAGGGAGACTAAATTAGGAAAAACGAG	3126						
QY	2027	GCTACGTGACCGGCGGCGCGCAGAGATCGTGAAGCTTGACCGGAGACCAACCAACAGA	2086						


```
QY 1487 TGCCTGAGTGTGCAAGCTGCTGCGCGCCCAAGCGCCTGACCGACATCGTGCCCTGA 1546
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3376 TAAGGCAATTATGTAACCTCTTAGAGAAACCAAGACACTAACAGAAAGTAATACCACTAA 3435

QY 1547 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGCGAGATCTCTGCGGAGCCGCTGCACG 1606
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3436 CAGAAGAGCAGAGCTAGAACTGGCAGAAAACAGAGAGATTCTAAAAGAACCACTATACG 3495

QY 1607 GCGTGTACTAGACCCAGCAAGGACCTGTGTGGCCGAGATCCAGAACGAGCCACACACC 1666
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3496 GAGTGTATTAGCCCATCAAGACTTATAGCAGAAATACAGAGCAGGGGCAAGGCC 3555

QY 1667 AGTGACCTACACAGATACAGGAGCCCTTCAAGAAACCTGAAGACCGGCAAGTACGCCA 1726
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3556 AATGGACATATCAAAATTTATCAAGAGCCATTAAATCTGAAAACAGGAAATATGCAA 3615

QY 1727 AGATGCCAGCGCCCAACACACACGACGTGAAGCAGCTGACGAGCGCGTGAGAAATCG 1786
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3616 GAATGAGGGGTGCCACACACTAATGATGTAACAAATTAACAGAGGCGAGTGCAAAATAA 3675

QY 1787 CCATGGAGAGCATCGTATCTGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAGG 1846
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3676 CCACAGAAAGCATAGTAATATGGGAAGACTCTCTAAATTTAAACTGCCCATACAAAGG 3735

QY 1847 AGACTCGGAGACTGCTGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1906
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3736 AAACATGGGAACATGCTGGACAGAGTATTGGCAAGCCACCTGGATTCTCAGTGGGAGT 3795

QY 1907 TCGTGAACACCCCCCTCGTGAAGCTGTGGTACAGCTGGAGAAAGAGCCCATATCG 1966
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3796 TTGTTAATACCCCTCCCTTATGTAAGAAATATGTTACCAAGTTAGAGAAAGAACCCATAG 3855

QY 1967 GCGCCGAGACTCTACGCTGGAGCGCCGCAACCGCAGACCAAGATCGGCAAGCGC 2026
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3856 GAGCAGAACTTCTATGTATAGTGGGCGAGCTACAGGGAGACTAAATTAGGAAAGCAG 3915

QY 2027 GCTACGTTACCGACCGGGCGGCGAGAGATCGTTGACCTGACCGAGACCAACCAACAGA 2086
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3916 GATATGTTACTATAGAGGAAGACAAAAGTTGTCAACCTTAACCTGACACAAACATCAGA 3975

QY 2087 AGACCGAGCTGCAGGCATCCAGCTGCGCTGCGAGACAGCGGAGCGAGGTGAACATCG 2146
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3976 AGATGAGTTACAAAGCAATTTATCTAGCTTTGCAGGATTCGGGATTAAGAAATAAACATAG 4035

QY 2147 TGACCCGACAGCCAGTACGCGCTGGGCAATCATCCAGGCCCGACACAGAGCGAGAGCG 2206
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4036 TAACAGACTCACATATGCAATTAGGAATCATTTAAGCACAAACAGATCAAGTGAATCAG 4095

QY 2207 AGCTGTGAACCGAGATCATCGACAGCTGATCAAGAAAGGAGAGGTGTACCTGAGCTGGG 2266
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4096 AGTTAGTCAATCAAAATATAGAGCAGTTAATAAAAAAGGAAAGGTCTATCTGCGATGGG 4155

QY 2267 TGCCCCCCCACAGGGCATCGCGGCAACAGCAGATCGACAGCTGGTGAAGCAAGGGCA 2326
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4156 TACCAGCACACAAAGGAATTTGAGGAAATCAACAAAGTAAATATAGTCAGTGTGGA 4215

QY 2327 TCAGCAAGTGTCTTCTCGGACGGATCGAT 2358
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4216 TCAGGAAGTACTATTTTAGATGGAATAGAT 4247
```

RESULT 11

```
us-09-943-286-3
; Sequence 3, Application us/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02.UT
; CURRENT APPLICATION NUMBER: US/09/943,286
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 8933
; TYPE: RNA
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)...(8933)
; OTHER INFORMATION: Sequence of transcripts produced from the BH10
; OTHER INFORMATION: plasmid.
us-09-943-286-3
```

```
Query Match 41.9%; Score 1029.2; DB 10; Length 8933;
Best Local Similarity 55.2%; Pred. No. 2,8e-161;
Matches 1329; Conservative 277; Mismatches 738; Indels 64; Gaps 7;

QY 14 TGGCCGAGGCGCATGAGCCAGG---CCACCAGCGCCACATCTGTGATGAGCGCAGCAACT 70
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1200 UGGUGAAGCAAUAGCGCAAGUAAACAAUACAGCUACCAUAAUGCAGAGAGGCAAU 1259

QY 71 TCRAAGGGCCCCAAGCGCATCATCAAGTCTCAACTGGGCAAGGAGGCCACATCGCCC 130
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1260 UAGGAACCAAGAAAGAUUGUUAAGUUUUAAGUGGCAAGAGGGGACACAGCCA 1319

QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGCAGGAGGCGCCACACAGA 190
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1320 GAAAUUGCAGGGCCCCUAGGAAGAGGCGUGUUGAAAUUGGAAAGAGGAGACACCAA 1379

QY 191 TGAAGGACTGACCGAGCGCCAGGCCAACTTCTTCCGCGAGGAGCTGCGCTTCCCCCAGG 250
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1380 UGAAGAAGUUGACUGAGAGACAGGCUAA-UUUUUUAGGGAAGAUUCUGGCCUUCUACAAG 1438

QY 251 GCAAGGCCCGCGAGT-----TCCCCAGCG 274
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1439 GGAAGGCCAGGAAUUUUUUCAGAGCAGACAGCCACAGCCCAACCAUUUUUUCAG 1498

QY 275 AGCAGAACCCGCGCCACACAGCCGCCAGCTGCAGGTGCGGG-----CGACA 328
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1499 AGCAGACAGAGCCACACAGCCGCCAGAGAGAGUUCUAGGUCUGGGUAGAGACAACA 1558

QY 329 ACCCCCGCAGGAGCGCGCGCGAGCCAGCCAGGCA-----CCCTGAACCTCCCCCAGA 382
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1559 ACUCCCCUUCAGAGCAGGAGCGGAGUAGACAAGGAACUGUAUCCUUUAACUUCUCUGA 1618

QY 383 TCACCTGTGTGACGCGCCCTGTGTGAGCTCAAGGTGGCGGCAGATCAAGAGAGGCC 442
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1619 UCACUUUUGGCAACGACACCCUUCGUCACAAUAAAGUAGGGGGCAACUAAAGGAAGCUC 1678

QY 443 TGCTGGACACCGCGCGCGACACACCCTGTGTGGAGGAGATGAGCCTGCCCGGCAAGTGA 502
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1679 UAUUAGAUACAGGAGCAGAUAGUACAGAUUUAAGAAAUAGUUUUGCCAGGAAGAUAGA 1738

QY 503 AGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCAAGGTGCGCCAGTACGACAGATCC 562
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1739 AACCAAAAUUAGUAGGGGAAUUGGAGGUUUUAUCAAGUAAAGACAGAUAGUAGAUAC 1798

QY 563 TGATCGAGATCTGCGGCAAGAGCCATCGCACCGTGTGATCGGCCCCACCCCGCTGA 622
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1799 UCAUAGAAAUUCUGUGGACAUAAAGCUAAGUACAGAUUUAUAGAGACCUACACCUGUCA 1858

QY 623 ACATCATCGGCGCGCAACATCTGTACCCAGCTGGGCTGCACCTTGAACCTTCCCCATCAGCC 682
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1859 ACAUAAUUGGAAGAAAUUCUGUACUCAGAUUGUUUGACUUUUAUUUUUUUUUUUUUAGCC 1918

QY 683 CCATCGAGACCGTCCCGTGAAGCTGAAGCTGAAGCCCTGACCGCCATCTGCGAGAGATGGAAGG 742
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1919 CUAUUGAGACUGUACCAAGUAAAUUUAAGCCAGGAUUGGAGGCCCAAAAGUUUUAACAAU 1978

QY 743 GGCCCTTGACCGGAGGAAGATCAAGCCCTGACCGCCATCTGCGAGAGATGCGAGAGG 802
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1979 GGCAUUGACAGAGAGAAAAUAAAGCAUUAUAGAAAAUUUUGUACAGAAAUUGAAAGG 2038

QY 803 AGGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTTACAACACCCCGCTGTTCGCCATCA 862
```


Db 1320 GAAAUUCAGGGCCCCUAGGAAAAAGGCGUCUUGGAAAAUUGGAAAGGAGGACACACAAA 1379
QY 191 TGAAGGACTGCACGAGCGCCAGCCCAACTTCTTCGCGAGGAGCTGSCCTTCCCCCAGG 250
Db 1380 UGAAGAUUUAUCUGACAGACAGCUGAA-UUUUUUAGGGAAGAUUCUGCCUUCUACAAG 1438
QY 251 GCAAGGCCCGCGAGT-----TCCCCAGCG 274
Db 1439 GGAAGGCCAGGAUUUUUUCUGAGCAGACCAGCCACAGCCCAACCAUUUUCUACG 1498
QY 275 AGCAAGACCGGGCAACAGAGCCCAACAGCGCGAGCTGCGAGTGCAGGTGCGCGG-----CGACA 328
Db 1499 AGCAGACAGAGCAACAGAGCCCAACAGAGAGAGCUGUAGAGGUGGUGAGAGACAACA 1558
QY 329 ACCCCGACGAGGCGCGCGCGAGCGCCAGGCA-----CCCTGAACCTTCCCCAGA 382
Db 1559 ACUCCCCUCAGAGCAGAGGCGAUAGACAGGAACUGUAUCCUUUAUCCUCCUAG 1618
QY 383 TCACCCCTGTGGCAGCGCCCTGCTGCTGAGCATCAAGTGGGCGCCAGATCAAGAGAGGCC 442
Db 1619 UCACUCUUGGCAACGACCCUUGUCACAAUAAAGAUAGGCGGCAACUAAAGAGACUC 1678
QY 443 TGCAGACACCGCGCGAGCAGCACCGTCTGGAGGAGATGAGCTCCCGCGCAAGTGA 502
Db 1679 UAUUAGAUACAGGAGAGAGUAGUACAGUUAUAGAAAUAGUUGGCCAGGAAGUAGA 1738
QY 503 AGCCCAAGATGATCGGCGGCGATCGGCGCTTCATCAAGTGGCGCAGTACGACCATCC 562
Db 1739 AACCANAUAUGAGGGGAUUGGAGGUUUUAUACAAGUAAGACAGUAGUACAGAUAC 1798
QY 563 TGATCGAGATCTGGGCAAGAGCCATCGGCACCGTGTGATCGGGCCCAACCCCGGTGA 622
Db 1799 UCATAGAAUCUGUGGACAUAAAGCAUAGUAGUACAUUAGUAGGACCUACACCUUCA 1858
QY 623 ACATCTGCGCCGCAACATCTGACCCAGCTGGGCTGCACCCCTGAACCTTCCCCATAGCC 682
Db 1859 ACAUAAUUGGAAGAAUUCUGUUGCCUGAGAUUGGUCACUUUAAAUUUUCCCAUAGCC 1918
QY 683 CCATCGAGACCGTGGCGTGAAGCTGAAGCCCGCGCATGGAGCGCCCAAGGTGAAGCAGT 742
Db 1919 CUUUGAGUGUACCAAGUAAAUAAAGCCAGGAAUGGAGGCCCCAANAAGUUAACAACU 1978
QY 743 GGGCCCTGACGAGGAGAGATCAAGCCCTGACCGCCATCTCGGAGGAGATGGAGAAG 802
Db 1979 GGCCAUUGACAGAGAAAAAUAAGCAUUAAGUAGAAAUUUGUACAGAAAUAGGAAAG 2038
QY 803 AGGCAAGATCACCAAGTCGCGCCGAGAACCCCTACAACACCCCGCTTCCGCCATCA 862
Db 2039 AAGGAAAAUUUCAAUUGGCGCCUGAGAAUCCACAUAUCCUAGAUUUUGCAUAA 2098
QY 863 AGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCA 922
Db 2099 AGAAAAAGACAGUACUAAUUGGAGAAAAUUAUAGAUUUUCAGAGACUUAUUAAGAGAA 2158
QY 923 CCAGAGCTTCTGGAGGTGACCTGGGCATCCCGCACCCCGCGCGCTGAAGAAGAAGA 982
Db 2159 CUCAGACAUUCUGGGAUUAUUAAGGAUACCAUCCCGCGAGGUGUAAAAAAGAAAA 2218
QY 983 AGAGCGTGAACGCTGGAGTGGCGGACGSCCTACTTTCAGCGTGCCTTGGACGAGACT 1042
Db 2219 AUCAGUACAGUUGGAGUUGGUGUAGUUAUUUUUUCAGUUCUUUAGUAGAGACU 2278
QY 1043 TCCGCAAGTACACCGCTTCCACATCCCAAGCATCAACACAGAGACCCCGGATCCGCT 1102
Db 2279 UCAGGAAGUUAUCGUAUUUACCAUACUAGUUAUAAACAUAGAGACACAGGAGUUAU 2338
QY 1103 ACAGTACAAAGTGTGCTGCCCGAGGTGGAAGGCGACCCCGCATCTTCCAGAGCAGCA 1162
Db 2339 AUCAGUACAUUGUGUUCACAGAGGUAUGGAAAGGAUACCAAGCAUUAUCCAAAGUAGA 2398
QY 1163 TGACCAAGATCTGAGCGCTTCCGCGCGCGCAACCCCGAGATCTGTACTACCA----- 1217
Db 2399 UGACAAAAAUUUUAGGCCUUUUUAAAAAACAACAAUCCAGACAUAGUUUUCUACAUAUA 2458

QY 1218 -GGCCCCCTGTAGCTGGCAGCGACCTGGAGATCGCGCAGCACCAGGCGCCAGATCGAGG 1276
Db 2459 UGGUAGUUAUUGUAGAGUAGUAGUAGAAUAGGCGAGCAGUAGAACAAAAUAGAGG 2518
QY 1277 AGCTGCGCAAGACCTGCTGCGTGGGCTTCACACCCCGGACAAAGACACAGAGG 1336
Db 2519 AGCUGAGACAAUCUCUUGAGUGGAGACUUACCAACACAGACAAAAACAUCAGAAAG 2578
QY 1337 AGCCCCCTTCTGCCCCAT-----CGAGTGTACCCCGGACAAAGTGGACCGTGCAGCCCA 1390
Db 2579 AACCUCAUUCUUUGAGUGGUUAUAGAACUCCAUCCUGAUAAUAGGACAGUACAGCCUA 2638
QY 1391 TCGAGCTGCGCGAAGAGAGCTGACCGCTGACGAGATCCAGAGCTGCTGGCGCAGC 1450
Db 2639 UAGUGUCGAGAAAAAGACAGCUGGAGUCUGAACAGUACAGAGUAGUAGGGGAAU 2698
QY 1451 TGAACCTGGCGCAGCAGATCTACCCCGCATCAAGTGGCGCAGCTGTGCAAGTGTGTC 1510
Db 2699 UGAUUGGCAAGUACAGUUAUACCCAGGAAUAAAGCAUUAUAGUAAACUCCUA 2758
QY 1511 GCGCGCCAGGCGCTGACCGACATCTGTCCTGACCGAGGAGCGCGAGTGTGAGTGG 1570
Db 2759 GAGGAACCAAGCACAACAGAGUAAUACCAUACAGAGAAAGACAGAGUAGAACUG 2818
QY 1571 CCGAGAACCGCGAGATCTGCGGAGCGCTGACGCGTGTACTACGACCCAGCAAG 1630
Db 2819 CAGANAACAGAGAUUUAAGAACCAAGUACUAGGAGUUAUAGUAAACUCCUA 2878
QY 1631 ACCTGTTGGCGAGATCCAGAGCGAGCGCACAGTGGACCTTACAGATCTACCAGG 1690
Db 2879 ACUUAUAGAGAAUACAGAGCGGCAAGGCAAGGCAUUAUAGGACAUUCAAUUAUUAAG 2938
QY 1691 AGCCCTTCAGAAACCTGAAGACCGCAAGTACGCCAGATGCGCACCGCCACACCAACG 1750
Db 2939 AGCAUUUAAAAUUCUGAAACAGGAAAUUAGCAAGUAGGGGUGGCCACACUAUAG 2998
QY 1751 ACCTGAAGCAGCTGACCGAGCGCTGCAGAGATCGCATGGAGAGCATGCTGTCTGGG 1810
Db 2999 AUGUAAACAAUUAACAGAGCGAGUGCAAAAUAAUACCAAGCAAGCAUAGUAAUAGG 3058
QY 1811 GCAGAGCCCCAAGTTCCGCTTGCCTTCCCATCCAGAGAGACCTGGGAGACCTGTGGACCG 1870
Db 3059 GAAGAGCUCUAAAUUUAUAAUACCCAUACAAAAGGAAACAUUGGAAACAUUGGAGAC 3118
QY 1871 ACTACTGGAGCGCACCTGGATCCCCAGTGGGAGTTCTGTGAACACACCCCGCTGTGA 1930
Db 3119 AGUAGGCGAAGCCACUGUAGUUCUGAGUGGAGUUUUUAUACCCCUUUAUGA 3178
QY 1931 AGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGCGCGAGACCTTCTACGTGGAGC 1990
Db 3179 AUUAGUAGCAGUUAAGAGAAAGACCCUAGUAGGAGCAGAAACCUUUAUAGUAG 3238
QY 1991 GCGCGCCAAACCGCGAGACCAAGATCGGCAAGCGCGCTACGTGACCGACCGGCGCGC 2050
Db 3239 GGGCAGCUAACAGGGAGACUAAUUAAGGAAAGCAGGAUUAUUAUAAAGAAAGAC 3298
QY 2051 AGAGATCTGAGCCTGACCGAGCACCAACAGAGACCGAGCTGCAGGCCATCCAGC 2110
Db 3299 AAAAGGUUCCUCCUUAACACAAACAAACAGAAACUGAGUUAACAGCAAGCAUUAU 3358
QY 2111 TGCCCTTACAGGACAGCGGAGGAGTGAACATCTGTGACCGAGACAGCAGTACGCCCTGG 2170
Db 3359 UACUUUGCAGGAUUCAGGAUUAAGAGUAAACAUUAUACAGACUCACAUUAGUAG 3418
QY 2171 GCATCATCCAGGCCCCCGCAAGAGCGAGAGCGAGCTGGTGAACCGATCATCGAGC 2230
Db 3419 GAUAUCAUAAAGCACACCAAGUAAAGUAGUUAUUAUUAUUAUUAUUAUUAUUAU 3478
QY 2231 AGCTGATCAGAGGAGAGGTGCTACTGTAGCTGGGTGCGCGCCCGACAGGGCATCGCG 2290
Db 3479 AGUUAUAAAAAGGAAAGGUCUUCUGGCAUUGGUGGACAGCACACAAAGGAUUGGAG 3538


```
QY 2291 GCAACGACGACATCCACAAGCTGTTGAGCAGGCGCATCCGCAAGGTGCTGTTCTCTGACG 2350
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 3539 GAAAGAACAGUAGUAAUUAUGACUGUGGAAUCAGGAAAUACUUAUUUUAUGAUG 3598
: || || || || || || || || || || || || || || || || || || || || || || ||

QY 2351 GCATCGAT 2358
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 3599 GAAUAGAU 3606
: || || || || || || || || || || || || || || || || || || || || || || ||

RESULT 13
US-09-943-286-9
; Sequence 9, Application US/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02.UT
; CURRENT APPLICATION NUMBER: US/09/943,286
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 8933
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the IAC-Bscr pseudo target
; NAME/KEY: mutation
; LOCATION: (4140)...(4159)
; OTHER INFORMATION: Mutated positions: 4140-42, 4145-47, 4152,
; OTHER INFORMATION: 4156-57, 4159
US-09-943-286-9

Query Match 41.9%; Score 1029.2; DB 10; Length 8933;
Best Local Similarity 55.2%; Pred. No. 2.8e-161;
Matches 1329; Conservative 277; Mismatches 738; Indels 64; Gaps 7;

QY 14 TGGCGGAGCCATGAGCCAGG---CCACCAGCGCCACATCTTGATGCGAGCGCAGCAACT 70
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1200 UGGUGAAGCAUAGGCAAGUACAAUACAGCUACCAUUAUGAUGCAGAGAGGCGAAU 1259

QY 71 TCAAGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGAGGAGGCGGCACATCGCC 130
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1260 UUAGGAACAAAGAAAGAUUGUUAUGUUAUGUUAUGUUAUGUUAUGUUAUGUUAUG 1319

QY 131 GCAACTGCCGCGCCCGCCGCAAGAGGCTGTGGAAGTGGCGCAAGGAGGCGCACAGA 190
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1320 GAAUUGCAGGCGCCCUAGGAAAGAGGCGUGUUGGAAUUGUGGAAAGAGGACACCAA 1379

QY 191 TGAAGGACTGCACGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCCCTTCCCGCAGG 250
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1380 UGAAGAUAUUGACUGAGAGACAGCGCUAA-UUUUUUAGGGAAGAUUCUGGCCUUCUACAAG 1438

QY 251 GCAAGGCGCCGAGT-----TCCCGAGCG 274
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1439 GGAAGCGCAGGGAUUUUUUCUAGAGCAGACACAGAGCAACAGCCCAACCAUUUUCUACG 1498

QY 275 AGCAGAACCCGCGCAACAGCGCCACAGCGCGAGCTGCGAGTGCAGGTGCGCGG-----CGACA 328
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1499 AGCAGACAGCGCCACAGCGCCCAACAGAGAGAGCUUCAGGUCUGGUGUAGAGACAACA 1558

QY 329 ACCCGCGCAGCGGCGCGCGCGCGCGCGCGCA-----CCCTGAACCTTCCCGCAGA 382
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1559 ACUCCCCUAGAGAGGAGGCGGAUAGACAGAGAAAGAAUUCUUUAUUCUCCUCAGA 1618

QY 383 TCACCTGTGGCAGCGCCCTCTGTTAGCATCAAGTGGCGGCGCGAGATCAAGAGGCC 442
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1619 UCACUCUUUGGCAACGACCGCCUGCACAUAUAAAGAUAGGGGCGCAACUAUAAAGAGCUC 1678

QY 443 TGCTGGACACCGCGCGAGCACACCTGCTGGAGGAGATGAGCCTGCCCGCGCAAGTGA 502
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1679 UAUUAGUACAGGAGCAGAUACAGUUAUAGAAAGAUUGUUUGCCAGGAAUGGA 1738
```

```
QY 503 AGCCCAAGATGATCGGGCGCATCGCGCGCTTATCAAGGTGCGCGCTACGACGATCC 562
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1739 AACCAAAAUAGUAGGGGAUUGAGGUUUUUAUCAAAGUAGACAGAUUGAUCAAGUAC 1798

QY 563 TGATCGAGATCTCGGGCAAGAACCATCGGCACCGTCTGATCGGCGCCCAACCCCGTGA 622
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1799 UCAUAGAAAUUGUGGACAUAAAGCUAUGUACAGUAUUAUAGUAGGACCUACACCCUGUCA 1858

QY 623 ACATCATCGCGCGCAACATGCTGACCCAGCTGGGCTGACCCCTGACTTCCCATCAGCC 682
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1859 ACAUAAUUGGAAGAAUUCUGUUGUCACUAGUUGUGCACUUUAAUUUCCCAUUAAGCC 1918

QY 683 CCATCGAGACCGTCCCGTGAAGCTGAAGCCGCGCATGAGCGCGCCCAAGGTGAAGCAGT 742
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1919 CUUUGAGACUGUACAGUAAAUUAAGCCAGAAUUGAUGGCGCCCAAAAGUUAACAAU 1978

QY 743 GGCCCTGACCGAGGAGAAATCAAGGCCCTGACCGCATCTGCGAGGAGATGAGGAAG 802
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 1979 GGCCAUUGACAGAAAGAAAUAAAGCAUUAUAGUAAAUUUGUACAGAAUUGGAAGG 2038

QY 803 AGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTACAAACACCCCGTGTTCGCCATCA 862
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2039 AAGGAAAUUUUCAAUAAUUGGCGCUGAGAAUCCAAUACAAUACUCCAGAUUUUGCCAUAA 2098

QY 863 AGAAGAGGAGCAGCACCAAGTGGCGCAAGCTGTGGACTTCCGCGAGCTGAACAGCGCA 922
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2099 AGAAGAGAGCAGUACUAAUUGGAGAAAUUAGUAAUUCAGAGAAUUAUUAAGAGAA 2158

QY 923 CCCAGGACTTCTGGGAGGTGCGAGTGGGCATCCCCACCCCGCGCGCTGAAGAAAGA 982
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2159 CUCAAGACUUCUGGGAAGUUAUUAUAGAAUACCAUCCCGCAGGCUUAAAAAGAAAA 2218

QY 983 AGACGCTGACCGTCTGAGCGTGGCGGCGCGCTACTTTCAGCGTGGCCCTGACGAGGACT 1042
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2219 AAUCAGAAUACAGUACUGGAGUGGUGAUGCAUUAUUUUCAGUUUCCCUUAGAAAGACU 2278

QY 1043 TCCCAAGTACACCGCTTACCATCCCGAGCATCAACAAGCAGAGACCCCGGCTCCGCT 1102
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2279 UCAGAGAUUAUUGCAUUAUACUACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2338

QY 1103 ACCAGTCAACAGTCTGCCCGCAGGCTGGAAGGCGAGCCCGCAGCATCTTTCAGAGCAGCA 1162
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2339 AUCAGUACAAUUGUUCUCCACAGGGAUGGAAGGAUACACAGCAUUAUUCAAAGUAGCA 2398

QY 1163 TGACCAAGATCTCGAGCGCTTCCCGCGCCCGCAACCCCGAGATCGTGATCTACCA----- 1217
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2399 UGACAAAUUUUAGAGCCUUUUUAAAAAUAUCCAGACAUAGUUUUAUUAUUAUUAUUA 2458

QY 1218 -GGCCCCCTGTAGCTGGGCGAGCGACCTGGAGATCGGCGCAGCAGCGCGCCCAAGATCGAGG 1276
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2459 UGGAUAGUUUUAUUGAUGAGUACUUGCUUAGAAUAGGCGCAGCAUAGAACAAAUAGAGG 2518

QY 1277 AGTGGCGCAGACCTGCTGGGCTGGGCTTTCACACCCCGCAGCAAGAACACAGAGG 1336
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2519 AGCUGAGACAACAUCUGUGAGGUGGAGCUUACACACAGCAACAAAAAUAUAGAGGAA 2578

QY 1337 AGCCCCCTTCTTCCCTCCAT-----CGAGCTGACCCCGCAGCAAGTGGACCTTGAGCCCA 1390
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2579 AACCUCAUCCUUUGAGGUGUUUAUGAACUCCCAUCCUGAUAAAUGGACAGUACAGCCUA 2638

QY 1391 TCGAGCTGCCCGAGAGGAGAGCTGGACCGTGAACGACATCCAGAGCTGGTGGGCAAGC 1450
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2639 UAGUGUGCCAGAGAAAGACAGUGGACUGCUUAUGACAUACAGAAUUAUAGUGGGGAAU 2698

QY 1451 TGAAGTGGCGCAGCGCATCTACCCCGGCATCAAGGTGCGCGAGCTGTGCAAGTGTGTC 1510
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2699 UGAUUGGGCAAGUCAGAUUUUACCCAGGGAUUAAGAAUAGGCAUUAUUAUAAUUCUUA 2758

QY 1511 GCGCGCCAGAGGCGCTGACCGACATCTGTCCTGACCGGAGGAGCGGAGCTGAGCTGG 1570
: || || || || || || || || || || || || || || || || || || || || || || ||
Db 2759 GAGGAACCAAGCACUAAACAGAGUAAUACCAUAAAGAAAGACAGAGACGUAAGCUGG 2818

QY 1571 CCGAGAACCGCGAGATCTCTGCGGAGCGCGCTGACGCGGTGTACTACGACCCAGCAAGG 1630
```

Db 2819 CAGAAAACAGAGAUUCUAAAAGAACCCAGUACAUAGGUGUAUUAUGACCCCAUCAAAAG 2878
QY 1631 ACCTGTTGGCGAGATCCAGAACAGGCGCCACAGCAGTGGAGCTACCAGATCACCAGG 1690
Db 2879 ACUUAUAGCAGAAUACAGACGAGCGGCAAGCCAAUGAGCAUAUCAAUUUAUCAAG 2938
QY 1691 AGCCCTTCAAGAACTGAAGACGGCAAGTACGCCAAGATGCGCACCGCCACACCAACG 1750
Db 2939 AGCAUUUAAAUCUGAAACAGGAAAUUUCAGAAUAGAGGGUGGCCACACUAUAG 2998
QY 1751 ACCTGAAGCAGCTGACCGGCGCTGAGAAAGATCCGCTGGAGAGCATGTCATCTGGG 1810
Db 2999 AUGUAAAACAUAUACAGAGCGCAGUGCAAAAAAUAAACACAGAAAGCAUAGUAUUGG 3058
QY 1811 GCAAGACCCCAAGTTCCGCTGCCATCCAGAGAGAGACCTGGGAGACCTGTGGACCG 1870
Db 3059 GAAAGACUCCUUAUUAACUACCCAUACAAGAAAGAAACAUAGGGAACAUUGGUGACAG 3118
QY 1871 ACTACTGGCAGGCACTGATCCCGAGTGGGAGTTCGTGAACACCCGCCCTCGTGA 1930
Db 3119 AGUAUUGGCAGCCACCGUAGUCCUGAGUGGAGUUUAUUAUACCCUCCUUAUGUGA 3178
QY 1931 AGCTGTGTTACAGCTGGAGAGAGGCCATCATCGCGCCGAGACCTTCTACGTGAGC 1990
Db 3179 AAUUAUGGUACCAAGUAGAGAAAGAACCCCAUAGUAGGAGCAGAAACCUUUAUAGUAG 3238
QY 1991 GCGCGCCCAACCGGAGACCAAGATCGGCAAGCGCGCTACGTGACCGACCGGGCGGC 2050
Db 3239 GGGCAGCUAACAGGGAGACUAAUUAUAGGAAGCAGGAUUAUUAACAAGGAAGAC 3298
QY 2051 AGAAGATCGTGAAGCTGACCGAGCACCAACACAGAGAGCCGAGCTGCGAGCCATCCAGC 2110
Db 3299 AAAAGGUUGUCCCUUAUAACACACAAAUACAGAAACUGAGUUAACAGAAUUAUC 3358
QY 2111 TGCCCTGTAGGAGAGCGAGGAGTGAACATCGTGACCGCAGCAGTACGCCCTGG 2170
Db 3359 UAGCUUUGCAGGAUUCAGGAUUAAGAAUUAUUAAGAAAGCAGGAUUAUUAACAUAUAG 3418
QY 2171 GCATCATCAGCGCCCGCAGACAGAGCAGAGCGAGCTGGTGAACCCAGATCATCGAGC 2230
Db 3419 GAUAUAUUAAGCAGACACACAGAAUUAAGUAGUACAGUAGUUAUUAUUAUAGAGC 3478
QY 2231 AGCTGATCAAGAGGAGAGGAGTGTACCTGAGCTGGGTGCGCGCCGACCAAGGCGATCGCG 2290
Db 3479 AGUUAUAAAAGGAAAGGUCUUCUGCAUGGUGUACCGACACACAGAAUUGGAG 3538
QY 2291 GCAACAGCAGATCGACAAGCTGGTGAAGCAAGGCGATCCGCAAGGTGCTGTTCTGTGAGC 2350
Db 3539 GAAUUAACAAGUAGUAAUUAUUAUAGUAGUGUGGAAUACAGGAAAUUAUUAUUAUAG 3598
QY 2351 GCATCGAT 2358
Db 3599 GAUAGAU 3606

RESULT 14
US-09-798-675-4
; Sequence 4, Application US/09798675
; Patent No. US20020106798A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: HIV VACCINES
; FILE REFERENCE: E056 2020
; CURRENT APPLICATION NUMBER: US/09/798, 675
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/186, 364
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/251, 083
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4

; LENGTH: 9544
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Construct of vaccine vector pGA2 and insert JS2 expressing cla
; OTHER INFORMATION: HIV-1 VL
; NAME/KEY: misc_feature
; LOCATION: (106)..(1641)
; OTHER INFORMATION: encoding inner core protein
; NAME/KEY: misc_feature
; LOCATION: (1401)..(3617)
; OTHER INFORMATION: encoding viral rproteins including reverse transcriptase
; NAME/KEY: misc_feature
; LOCATION: (3708)..(6334)
; OTHER INFORMATION: induces high level of virus particles
; NAME/KEY: misc_feature
; LOCATION: (3847)..(6518)
; OTHER INFORMATION: encodes Rev regulating transfer of RNA to cytoplasm
; NAME/KEY: misc_feature
; LOCATION: (4102)..(6660)
; OTHER INFORMATION: encoding ADA Env
; NAME/KEY: misc_feature
; LOCATION: (6672)..(9544)
; OTHER INFORMATION: Vaccine vector pGA2
US-09-798-675-4

Query Match 41.6%; Score 1022.8; DB 10; Length 9544;
Best Local Similarity 66.5%; Pred. No. 3.2e-160;
Matches 1602; Conservative 0; Mismatches 742; Indels 64; Caps 7;

QY 14 TGCCCGAGGCGCATGAGCAGG---CCACCAGCGCCACATCCTCTGATCAGCGCAGCAACT 70
Db 1193 TGCTGAAGCAATGAGCCAAAGTAACAAATACAGTACCATAATGATCAGAGAGGCAATT 1252
QY 71 TCAAGGGCCCCCAGCGCATCATCAAGTCTTCACTCGCGGAGGAGGCGCACATCGCC 130
Db 1253 TTAGGAACCAAGAAGATGTTAAGAGCTTCAATACGCGCAAGAGGCGCACAGCCCA 1312
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAAGGCTGCTGGAAGTGCAGGAGGCGGCGCCACAGA 190
Db 1313 GAAATTGACGGGCCCTTAGGAAAAGGCGCAGCTGGAAAAGCGGAAAGAGGAGGACACCAA 1372
QY 191 TGAAGGACTCACCGAGCGCCGACCACTTCTTCCGCGAGGACCTTGGCCTTCCCCCAGG 250
Db 1373 TGAAGAGATTCTAGAGAGACAGGCTAA-TTTTATGAGGAAGATCTGCGCTTCTCTACAAG 1431
QY 251 GCAAGGCCCGCGAGT-----TCCCCAGCG 274
Db 1432 GGAAGGCCAGGGAATTTTCTCAGAGCAGACACAGCCCAACAGCCCACTTTCTTTCAG 1491
QY 275 AGCAGAACCGCGCCACAGCCCGCAGCGCGAGCTGCGAGTGCAGG-----CGACA 328
Db 1492 AGCAGACAGCCACAGCCCGCCAGAGAGAGCTTCAAGTCTGGGTAGAGACACACA 1551
QY 329 ACCCCCGCAGCGCGCGCGCGCGAGCGCCAGGSCA-----CCTGAACCTTCCCCCAGA 382
Db 1552 ACTCCCTCTCAGAAGCAGGAGCGGATAGACAAGGAAGTGTATCTTTAACTTCCCTCAGA 1611
QY 383 TCACCTCTGTCAGCGCGCCCTTGGTGAAGATCAGGTTGGGGCGGCGCAGATCAAGAGGCC 442
Db 1612 TCACCTCTTGGCAACGACCCCTCTGTCACAATAAAGATAGGGGGGCAACTAAAGGAAGCTC 1671
QY 443 TGCTGGACACCGCGCGCGCAGACACCGTCTGGAGGAGATGAGGCTCGCGCGCAAGTGA 502
Db 1672 TATTAGATACAGGAGCAGATGATACATATTAGAAAGATGAGTTTCCAGGAGACATGGA 1731
QY 503 AGCCCAAGATGATCGCGCGGATCGCGGCTTCATCAAGGTGCGCCAGTACGACAGATCC 562
Db 1732 AACCAAAATGATAGGGGGAATTGGAGGTTTATCAAAAGTAAGACAGTATGATCAGATAC 1791
QY 563 TGATCAGATCTCGGCGCAAGAGGCCATCGGCACCGTGTGATCGGCGCCCGCCCGCTGA 622
Db 1792 TCATAGAAATCTGTGGACATAAAGCTATAGTACAGTATTAGTAGGACCTACACCTGTCA 1851

QY 623 ACATCATCGCGCGCAACATGCTGACCCAGCTGGGTGACCCCTGAACCTTCCCATCAGCC 682
DB 1852 ACATAATTTGAAGAAATCTGTTGACTCAGATGTTGACITTAATTTTCCCATAGCC 1911
QY 663 CCATCGACACCGTGGCGGTGAAGCTGAAGCCGGCATGAGCGGCCCCCAAGGTGAAGCACT 742
DB 1912 CTATTGAGAGCTGATACCACTAAATTTAAAGCCAGGAATGATGGCCCAAAAGTTAAACAA 1971
QY 743 GGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTGCGAGAGATGGAGAAGG 802
DB 1972 GGCCATTGACAGAGAGAAATTAAGACATTTAGTGAATTTGTACAGAAATGGAAAGG 2031
QY 803 AGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTACAAACACCCCGCTGTTGCGCATCA 862
DB 2032 AAGGAAATTTCAAAATTTGGCGCTGAGAAATCCATACAAATATCCACTATTTGGCATAA 2091
QY 863 AGAAGAGCAGACAGCAAGTGGCGCAAGCTGGTGGACTTCCGGAGCTGAACAGCGCA 922
DB 2092 AGAAGAGCAGACTACTAATGGAGAAATTTAGTAGATTTTCAGAGAACTTAATAGAGAA 2151
QY 923 CCCAGGACTTCTGGAGGTGCGAGCTGGGCATCCCCACCCCGCGCTGGAAGAGAAGA 982
DB 2152 CTCAGACTTCTGGAGTTCATTTAGATAGCAATCCACATCCCGCAGGGTTAAAGAGAAA 2211
QY 983 AGAGGTGACCGTGTGAGCTGGCGACGCTACTTACAGCGTCCCGCTGGACGAGGACT 1042
DB 2212 AATCAGTAACAGTACTGGATGGGTGATGCATATTTTCAGTTCCTTTAGTAGAAGACT 2271
QY 1043 TCCGCAAGTACACCGCTTCCACATCCCGACATCAACAGCAGACCCCGCGCTGAGAGAGCA 1162
DB 2332 ATCAGTACATGTCTTCCACAGGATGGAAGGATCACAGCAATATTCACAGTAGCA 2391
QY 1163 TGACCAAGATCTGGAGCCCTTCGCGCCCGCCCAACCCGAGATCGTGAATACCA ---- 1217
DB 2392 TGACAAAATCTTAGAGCCCTTTTAAAAAACAATAATCCAGACATAGTTATCTATCAATACA 2451
QY 1218 -GGCCCCCTCTAGCTGGGACGCGCTGGAGATCGCGCAGCACCCTGCGCAAGATCGAGG 1276
DB 2452 TGAACGATTTGTATGTAGGATCTGACTTAGAATAGGCGCATAGAACAAAATAGAGG 2511
QY 1277 AGCTGGCGAGCAGCTGCTGGCTGGGGCTTCAACCCCGCCGACAAAGACACAGAGG 1336
DB 2512 AGCTGAGACAACTCTGTTAGGTGGGACTTACCACACAGACAAAACATCAGAAAG 2571
QY 1337 AGCCCTCTCTCTGCCAT-----CGAGCTGACCCCGACAGTGGACCGTGAGCCCA 1390
DB 2572 AACCTCATCTCTTGGATGGGTTATGAATCCATCCCTGATATTAATGGACAGTACAGCCTA 2631
QY 1391 TCGAGCTGCCCGAGAGAGAGCTGGACCGTGAACGACATCCAGAACTGTTGGGCAAGC 1450
DB 2632 TAGTGCTGCCAGAAAAGACAGCTGGACTGTCAATGACATACAGAACTTAGTGGGGAAT 2691
QY 1451 TGAATCGGCGAGCAGATCTACCCCGGCATCAAGGTGCGCGAGCTGTGAAGCTGTGC 1510
DB 2692 TGAATCCGAAGTCAAGATTTACCCAGGGATTAAGTAAAGCAATTTATGTAACCTCTTA 2751
QY 1511 GCGCGCCAGAGCCCTGACCGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1570
DB 2752 GAGAACCAAGGACCTTACAGAGTAAATACCACTTAACAGAGAGAGCAGAGTAACTGG 2811
QY 1571 CCGAGACCCGAGATCTCTCGCGAGCGCTGCGAGCGGTGCTAGCAGCCCGCCAGCAGG 1630
DB 2812 CAGAAACAGAGAGATCTTAAGAGACCACTAGTACATGAGTGTATTATGACCCATCAAG 2871
QY 1631 ACCTGGTGGCGAGATCCAGAGCAGGCGCCAGCAGGAGTGGACCTACAGATCTACAGG 1690
DB 2872 ACTTAATAGCAGAAATACAGAGCAGGCGGCAAGGCCAATGGACATATCAATTTATCAAG 2931

QY 1691 AGCCCTTCAAGAACCTGGAAGCCGCAAGCTAGCCNAGATGCGCACCCGCGCCACCAAGC 1750
DB 2932 AGCCATTTAAATCTGAAAAAGGAAATATGCAAGATGAGGGGTGCCCCACACTAATG 2991
QY 1751 AGCTGAAGCAGCTGACCCGAGGCGCTGCAGAAATGCCATGGAGAGCATCTGTGATCTGG 1810
DB 2992 ATGTAACAAATTAACAGAGGCGAGTCAAAATAACACAGAAAGCATAGTAATATGGG 3051
QY 1811 GCAAGACCCCAAGTTCCTCCCTGCGCATCCAGAGAGACCTGGGAGACCTGGTGGACG 1870
DB 3052 GAAGAGCTCTTAATTTAACTACCATCAAAAGGAAACATGGGAAACATGGTGACAG 3111
QY 1871 ACTACTGCGAGCGCACTGGATCCCGAGTGGAGTTCGTGAACACCCCGCCCTGGTGA 1930
DB 3112 AGTATTGCAAGCCACTGGATTCCTGAGTGGAGTTTGTAAATACCCCTCTCTTAGTGA 3171
QY 1931 AGCTGTGTACAGCTGGAGAGAGCCCATCATCGCGCCCGAGACCTTCTACGTGGAG 1990
DB 3172 AATTATGTTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGTG 3231
QY 1991 GCGCGCCCAACCGCAGAGACCAAGATCGGCAAGCCCGCTACGTGACCGCGGCGCGC 2050
DB 3232 GGGCAGCTTAACAGGAGAGACTAAATTTAGGAAAAGCAGGATATGTTACTAACAAAGGAAG 3291
QY 2051 AGAAGATCTGAGCGCTGACCGAGACCAACCAAGACCGAGCTGCGAGCCCATCCAGC 2110
DB 3292 AAGAGTTGTTCCCTTAACATACACAAATCAGAAACTCAGTTACAAAGCAATTTATC 3351
QY 2111 TGGCCCTCGAGAGACGCGGAGCTGAACATCGTACCGACAGCAGCTACGCCCTGG 2170
DB 3352 TAGCTTTGAGGATTCAGGATTAGAAAGTAAACATAGTAACAGACTCAACAATGATAG 3411
QY 2171 GCATCATCCAGGCGCCGCGCAGAGAGCGAGAGCGAGCTGGTGAACAGATCATCGAGC 2230
DB 3412 GAATCTCAAGCAGACACCAAGTAAAGTAAATCAAGTTAGTCAATCAAAATATAGAGC 3471
QY 2231 AGCTGATCAAGAAAGAGAGGTGTACCTGAGCTGGGTGCGCCCGCCACAAAGGCGCGG 2290
DB 3472 AGTTAATAAAGAAAGAAAGGTCTATCTGGCATGGGTACCAGCAGCAAAAGGAATGGAG 3531
QY 2291 GCACGAGCAGATCGACAAAGCTGGTGAAGCAAGGCGATCCGCAAGGTGTGTTCTCGAGC 2350
DB 3532 GAAATGAACAAGTAGATAAATTAGTCAAGTGGTAATCAGGAAATACTATTTTTAGATG 3591
QY 2351 GCATCGAT 2358
DB 3592 GAATAGAT 3599

RESULT 15
US-09-798-675-5
; Sequence 5, Application US/09798675
; Patent No. US20020106798A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: HIV VACCINES
; FILE REFERENCE: E056 2020
; CURRENT APPLICATION NUMBER: US/09/798,675
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/186,364
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/251,083
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 9918
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Construct of vaccine vector pGal and vaccine insert expressing
; OTHER INFORMATION: ade B HIV-1 Gag-Po
; NAME/KEY: misc_feature

LOCATION: (106)..(1641)
OTHER INFORMATION: encodes proteins of viral inner core
NAME/KEY: misc_feature
LOCATION: (1401)..(3617)
OTHER INFORMATION: encodes viral proteins but not integrase
NAME/KEY: misc_feature
LOCATION: (3708)..(5715)
OTHER INFORMATION: regulates high-level production of HIV genes
NAME/KEY: misc_feature
LOCATION: (3847)..(5944)
OTHER INFORMATION: encodes protein regulating the transfer of RNA to cytoplasm
NAME/KEY: misc_feature
LOCATION: (3939)..(4181)
OTHER INFORMATION: gene participates in viral assembly and budding
NAME/KEY: misc_feature
LOCATION: (4099)..(4941)
OTHER INFORMATION: encodes truncated form of viral coat protein
NAME/KEY: misc_feature
LOCATION: (6098)..(9918)
OTHER INFORMATION: vaccine vector pGAL

US-09-798-675-5

Query Match 41.6%; Score 1022.8; DB 10; Length 9918;
Best Local Similarity 66.5%; Pred. No. 3.2e-160;
Matches 1602; Conservative 0; Mismatches 742; Indels 64; Gaps 7;
QY 14 TGGCCGAGGCGATGAGCCAGG---CCACCAGCGCCACATCTCTGATCGCAGCGCAGCAACT 70
DB 1193 TGGCTGAAGCAATGAGCCAGTAGTACAAATACAGCTACCAATGATCGAGAGGCGCAATT 1252
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGCCTTCAACTCGCGCAGGAGGCGCCATCGCC 130
DB 1253 TTAGGAACCAAGAAAGATGGTAAAGAGCTTCAATACGCGCAAGAGGCGCACACGCCA 1312
QY 131 GCAACTCGCGCGCCCGCCGCAAGAGGCTGCTGGAGTGGCGCAAGAGGCGCCACACAGA 190
DB 1313 GAAATTCAGGGCGCCCTAGGAAAAGCGAGCTGGAAGAGCGGAAAGAGAGACACAAA 1372
QY 191 TGAAGAGTCAAGCGAGCGGCGAGCCCAACTTCTTCCGCGAGGAGCTTCCGCCAGG 250
DB 1373 TGAAGATTGTACGTGAGAGACAGGCTAA-TTTTATTAGGAGAGTCTGGCCTTCTTACAAG 1431
QY 251 GCAAGGCGCGGAGT-----TCCCCAGCG 274
DB 1432 GGAAGGCGAGGAATTTTCTTCAGAGCAGACAGAGCCACAGCCCGCCACCATTTCTTCAG 1491
QY 275 AGCAGACCGCGCCCAAGCGCCCGCCAGCGCTGCGAGCTGCGCGG-----CGACA 328
DB 1492 AGCAGACCGAGCCCAAGCGCCCGCCAGAGAGCTTCAAGTCTGGGGTATAGAGACACA 1551
QY 329 ACCCGCGAGCG 382
DB 1552 ACTCCCGCTCAGAGCAGGAGCGCGATAGACAAGAACTGTATCCTTTAACTTCCCTCAGA 1611
QY 383 TCACCCCTGTGCGAGCGCCCGCTGTGAGCATCAAGGTGGCGCGCGCGCGCGCGCGCGCGCG 442
DB 1612 TCACTCTTTGGCAACGACCGCTCTGCAATAAAGATAGGGGGCACTAAAGAAAGCTC 1671
QY 443 TGCTGGACACCG 502
DB 1672 TATTAGATACAGGAGCAGATGATACAGTATTAGAGAAATGAGTTTGGCCGAGAGATGGA 1731
QY 503 AGCCCAAGATGATCGGGCGCTCGGGCGCTTCATCAAGGTGCGCGCGCGCGCGCGCGCGCG 562
DB 1732 AACCAAAATGATAGGGGAATTGGAGTTTATCAAGTAAGACAGTATGATCAGATAC 1791
QY 563 TGATCGAGATCGCGGCAAGAGCGCATTCGCGACCGTGTGATCGCGCGCGCGCGCGCGCG 622
DB 1792 TCATAGAAATCTGTGGACATAAAGCTATAGGTACAGTATTAGTAGGAGCTACAGCTGTCA 1851
QY 623 ACATCATCG 682
DB 1852 ACATAATTGGAAGAAATCTGTGTGACTCAGATTGGTTGCTTAAATTTTCCCAATFAGCC 1911

QY 683 CCATCGAGACCGTCCCGTGAAGCTGAAGCCCGCATGGACGCGCCCAAGGTGAAGCAGT 742
DB 1912 CTATTGAGACTGTACCAGTAAATTAAGCCAGGAATGGATGCCCAAAAGTTTAAACAAT 1971
QY 743 GGCCCTGTACCGAGGAGAGATCAAGCCCTGACCCGCTATCTCGAGAGAGATGAGAGAG 802
DB 1972 GGCCATTGACAGAAGAAAAAATAAAGCATTAGTAGAAATTTGTACAGAAATGGAAGG 2031
QY 803 AGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTACAACACCCCGCTGCGCCATCA 862
DB 2032 AAGGGAATAATTTCAAAATTTGGCCCTGAGAAATCCATACTTCCAGTATTTGCCATAA 2091
QY 863 AGAAGAGGAGCAGCACCAAGTGGCGCAAGCTGTGGACTTCCGCGAGCTGAACAAGCGCA 922
DB 2092 AGAAAAAGACAGTACTAATGAGAAAAATAGTAGATTTCAGAGAACTTAAATAAGAGAA 2151
QY 923 CCCAGGACTTCTGGGAGGTGCAGCTGGGCATGCCCGACCCCGCGCGCTGAAGAGAGAA 982
DB 2152 CTCAGAGACTTCTGGGAAGTTCAATTTAGGAATACCACTATCCCGCAGGTTTAAAAAGAAA 2211
QY 983 AGACGCTGACCGTCTGGAGCTGGCGCAGCGCTACTTTCAGCGTCCCGCTGACGAGACT 1042
DB 2212 AATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTTCAGTTCCCTTAGATGAAGACT 2271
QY 1043 TCCGCAAGTACACCGCTTCCACCATCCCGCAGCATCAACAAGAGAGACCCCGCGCATCCGCT 1102
DB 2272 TCAGGAAGTATATGCATTTACCATACCTAGTATAACAATAGACACACCGAGGATTAGAT 2331
QY 1103 ACCAGTACACAGTCTGCCCGAGGCTGGAGGCGCAGCCCGCGCTGTAAGAGAGAA 1162
DB 2332 ATCAGTACAATGTCTTCCACAGGATGGAAGGATCACCAGCAATATTTCAAAATAGCA 2391
QY 1163 TGACCAAGATCTCGAGCGCTTCCCGCCCGCAACCCCGAGATCGTGATCTACCA----- 1217
DB 2392 TGACAAAAATCTTAGAGCTTTTAAAAAACAATAATCCAGACATAGTTATCTATCAATACA 2451
QY 1218 -GGCCCCCTCTAGCTGGGCGAGCGCTGGAGATCGGCGCAGCACCGCGCGCTGAGAGG 1276
DB 2452 TGAACGATTTCTATGTAGGATCTGACTTAGAAATAGGCGCAGCATAGAACAAAAATAGAG 2511
QY 1277 AGCTGGCGCAGCACCTGCTGGCGTGGGCTTCCACCCCGCGAGAGAGCAGCAGAGAGG 1336
DB 2512 AGCTGAGACAACTCTGTTGAGGTGGGAGCTTCCACACCGAGCAAAAAAATCAGAAAG 2571
QY 1337 AGCCCCCTTCTGCCCCAT-----CGAGCTGCACCCCGCAGAAAGTGGACCTGCGAGCCCA 1390
DB 2572 AACCTCCATTCTTGGTGGTTATGAACCTCCATCTCTGATAAATGGACAGTACAGCCTA 2631
QY 1391 TCAGAGTGGCGGAGAGAGAGTGGCCGTGAGAGCATCCAGAGCTGAGAGCTGGTGGCAGC 1450
DB 2632 TAGTGTCTGCCAGAAAAAGACAGCTGGAGTCTCAATGACATACAGAAAGTTAGTGGGAAAT 2691
QY 1451 TGAAGTGGCGCAGCAGATCTACCCCGCATCAAGGTGCGCGAGCTGCAAGCTGTGCTGC 1510
DB 2692 TGAATCCGCAAGTCAGATTTTACCCAGGATTAAGTAGGCAATTTATGAATCTCTTA 2751
QY 1511 CGCGCGCAAGCGCTGACCGACATCTGCGCCCTGACCGAGGAGCGCGAGCTGGAGCTGG 1570
DB 2752 GAGGAACCAAGCAGTACAGAAATTAACCACTAACAGAGAGAGCAGAGCTAGAACTGG 2811
QY 1571 CCGAGAACCGGAGATCTGCGCGAGCGCGTGGAGCGGTGTACTACACCCCGCAGCAGG 1630
DB 2812 CAGAAAAACAGAGATTTCTAAAGAACCCAGTACATGGAGTGTATTATACCCCATCAAAAG 2871
QY 1631 ACCTGGTGGCGAGATCCAGAGCAGCGCCAGCACCTGGACCTTACCAGATCTACCAGG 1690
DB 2872 ACTTAATAGCAGAAATACAGAGAGCGGCGAGGCGCAATGGACATATCAATTTATCAG 2931
QY 1691 AGCCCTTCAAGAACTGAAGACCGCGCAAGTAGCGCAAGATCGCGACCCCGCCACACCAAG 1750
DB 2932 AGCCATTTAAANCTGAAACAGAGAAATATGCAAGAAATGAGGGGTGCGCCACACTAATG 2991

1751	QY	ACGTGAGCAGCTGACCGAGCGCGTGCAGAAGATGCGCAATGGAGAGCATCGTATCTGGG	1811
2992	Db	ATGTAAAACAATTACAGAGCAGCTGCAAAAAATTAACACAGAAAGCATAGTAATATGGG	3051
1811	QY	GCAAGACCCCAAGTTCCGCGTCCGCATCCAGAGGAGACCTGGGAGACCTGGTGGACCG	1870
3052	Db	GAAGACCTCTAAATTAAACTACCCATACAAAGGAAGAACATGGGAACATGGTGACAG	3111
1871	QY	ACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTGTTGAACACCCCCCCTCGTGA	1930
3112	Db	AGTATTGGCAAGCCACCTGGATTCTCGTGGGAGTTTGTTAATACCCCTCCTTTAGTGA	3171
1931	QY	AGCTCTGTATCCAGCTGCGAAGGAGGCCATCATCGCGCGGAGACCTTCTAGCTGGACG	1990
3172	Db	AAATTATGTATCCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATG	3231
1991	QY	GGCGCCCAACGGGAGACCAAGATCGGCAAGCGCGGTACTGTTGACCGACCGGGCGCGC	2050
3232	Db	GGCGAGCTTAACAGGAGACTAAATTAGAAAAGCAGGATATGTTACTACAAGGAAGAC	3291
2051	QY	AGAAGATCGTGAGCGTGACCGAGACCAACACAGAGACCGAGCTGCAGGCCATCCAGC	2110
3292	Db	AAAAGGTTGTCCCGCTAACTAACACAAATCAGAAAACCTCAGTTTACAAGCAATTTATC	3351
2111	QY	TGGCCCTCGCAGACAGCGGCAGCGAGTGTAACATCTGTGACCCGACACCCAGTACGCCCTGG	2170
3352	Db	TAGCTTTCCAGGATTGAGGATTAGAAGTAACATAGTAACAGACTCACAATATGCATTAG	3411
2171	QY	GCATCATCCAGGCCCGCCGACAAAGAGCGAGACGAGCTGGTGTGAACACAGATCATCGAC	2230
3412	Db	GAATCATTTCAAGCACACACAGATAAAGTCAATCAGAGTTAGTCAATCAAAATAATAGAC	3471
2231	QY	AGCTGATCAAGAAGAGAGAGGTGTACTGTAGCTGGGTGGCCGCCACAAAGGGCATCGGGC	2290
3472	Db	AGTTAATAAAAAGGAAAGGTCTATCTGCATGGGTATCCAGCACACAAAGGAATGGAG	3531
2291	QY	GCAAGCAGCATGCACAAGCTGTTGAGCAAGGSCATCCGCAAGGTGCTGTTCTCTGGACG	2350
3532	Db	GAATGACACAGTAGATAAATTAGTCAGTCTGGAATCAGGAAAATACATATTTTAGATG	3591
2351	QY	GCATCGAT	2358
3592	Db	GAATAGAT	3599

Search completed: February 10, 2003, 21:12:28
Job time : 154.404 secs

